

1 TGCCCGCTGC CCGCCCGCAG TTCCCGGCC CGCTGGCCCC AGTCATGGCG  
51 AAGCAGTACG ATGTGCTGTT CCGGCTGCTG CTGATCGGGG ACTCCGGGGT  
101 GGGCAAGACC TGCCTGCTGT GCCGCTTCAC CGACAACGAG TTCCACTCCT  
151 CGCACATCTC CACCATCGGT GTTGACTTTA AGATGAAGAC CATAGAGGTA  
201 GACGGCATCA AAGTGCGGAT ACAGATCTGG GACACTGCAG GGCAGGAGAG  
251 ATACCAGACC ATCACAAAGC AGTACTATCG GCGGGCCCAG GGGATATTTT  
301 TGGTCTATGA CATTAGCAGC GAGCGCTCTT ACCAGCACAT CATGAAGTGG  
351 GTCAGTGACG TGGATGAGTA CGCACCAGAA GGCGTCCAGA AGATCCTTAT  
401 TGGGAATAAG GCTGATGAGG AGCAGAAACG GCAGGTGGGA AGAGAGCAAG  
451 GGCAGCAGCT GCGGAAGGAG TATGGCATGG ACTTCTATGA AACAAAGTGCC  
501 TGCACCAACC TCAACATTAA AGAGTCATTG ACGCGTCTGA CAGAGCTGGT  
551 GCTGCAGGCC CATAGGAAGG AGCTGGAAGG CCTCCGGATG CGTGCCAGCA  
601 ATGAGTTGGC ACTGGCAGAG CTGGAGGAGG AGGAGGGCAA ACCCGAGGGC  
651 CCAGCGAACT CTTCGAAAC CTGCTGGTGC TGAGTCTGT GTGGGGCACC  
701 CCACACGACA CCCCTCTTCC CTCAGGAGG CCGTGGGCAG ACAGGGGAGC  
751 CCGGGCTTTG CCCTGTGCT GTCTCTCGT GTGATGACCC TATTGAGTAT  
801 CAGTAGCCAC TACTCCCCCT GCCTGGCCCT GAGAGCGGCT CTGCTGTCAT  
851 CTCAAGCAGC CCCTGTCCCC AGCCCGTCCA CCCTGGAGTG GTCTTCTTCA  
901 GCCTGTTTCC CCAGCCACAG GCCTGCTACG ACCCCACGA TGTGCCGCAA  
951 GCACTGTCTC ACCATCCCGC ACCCACCAGA CAACAGCCAG GGCTGGAGTC  
1001 CAGGCCACTT TCAGCTGCTC CTTTCTCCGT GCATCGTGTC TCTTCTCTGC  
1051 TTTTCTCTC TTCCCCACT TCTTCTTCTC TGACCCCTCC CCTCCGGTGC  
1101 GTTTCGTATC AAAGCTCCTC AAACCCCGTC CCCCCTGTGT CCTGCTGTGT  
1151 GCAGCTCGCT CTTTCTTCC TTCTTAAGCT ATCCAAGGGG ATGGACCCAG  
1201 GCTCGTGGGG AGGTTCCACC CTTGGATCCA GGAAGAACCC TCCACCCTGC  
1251 CTCGTGGGTG GGCCAAAGGC TACAGGGTGC TTCTTCCTCT TCCCCACCC  
1301 CCACTGTCCC TCATGTGCCA TGGGCCTGCC TCCCAGTGA CTGCGAAAG  
1351 TGGAGCATCG AGGTAGGAGG GAAACAGCAA CCGGGGAGTC CTCGAGCCTG  
1401 GGGCTGCCCT ACCTCTACCC ATTCCCCGAC CAGAGCTTGT CCCTTGCTTG  
1451 GCTGCCCCGC TGCTCTTTG GGGAACTGAG CTCAGAGGCA GGTGCTTCAG  
1501 AGAAGGAAAC AAAATGAGGG GTGGCAGGGA TAAAAAGTCA CCTCCATTCT  
1551 CTACCTCCCA TGCAGCATGA ACACAATTTT TCTCCACCTG GCTCCCAAAT  
1601 TTAAAGATGT GGACCAAGGC CTGTGGGTAC TCCAGGGGCA AGGAGAGCCC  
1651 TGGGGTCAGT GACACTGTCA GGCCAACCAT GCACTCCACA AAGGGGAGCA  
1701 TTTGGAAATG AAGGACTAGC TCCTATGTAT CAGGTTAAGA GCAAGGGAGA  
1751 GCTGGCCAGG GACAGCAGTT TGCACAGCAG AGGGGAATGT AGCAACAGCA  
1801 GGGCCTCCTA GGCCCCATCT TCCATTCTTT AGGTAAGAAG AGCATTTCTT  
1851 CAGACTCCCA GGCGGAGGAC TGAGCCTAGC CTTAGCAAC CAAGGTTCTC  
1901 CTGGGACCCA AAGTTTATGG GAGAAGGGCA AAGACTTCAT GGGAGAGAG  
1951 AAGGAAGGCC CTGGGTAGAA ACGCTTGGTG CTGTTCTCTT TGGCCTTTAA  
2001 GACAAAGCGC TCATCTTGCC CTCTACCTCC TGATAGGCTT GAGGTTTGC  
2051 CAACCACACT GTGGCTACAG GTGGAGGGAA GAGGACTCCT TCCTCCAGAG  
2101 TGCTATGTTT AGGAAGTTT TTTAACCCCA TATGGCCCAA GAGTAGCTCG  
2151 TAGGAGGCCC TTTAAAGACG GAACAAGTAA TTTACCAGTT CTACTGGGGT  
2201 TCCTGCCCCA CGTCCCAAGG TGGGCGAGGC CTAGGAAGAG GGTCACTCTT  
2251 AAGCCACACA TTAGCTGCAC TGCGTGGCTG CAGCCAAAAC AAAGAACTGG  
2301 GTGTTGAGTA TTCATCAACT AAGAACCAA ATCCAGGGCA CTCATATGTG  
2351 AAGGATAAGA ACCTCACTTC CTTACTCCTC CAAAAAGAAG TGGGGAAAGA  
2401 ACCATCAAAC CTTTCTCTCT GACTTACCAA ACCAGGAAA CAGCAGGAGA  
2451 GGGTGGCTCA GGACTTAGGG ACAGGGTATA GCTTAGATGG TGGAAGCAA  
2501 AGGAGAGCAG GAAGTTGTAA ATCACTGGCT AATGAGAAAA GGAGACAGCT  
2551 AACTCTAGGA TGAAGCTGTG ACTAGGCTGG AGTTGCTTCC TTGAAGATGG  
2601 GACTCCTTGG GTATCAAGAC CTATGCCACA TCACACTGGG GCTAGGGAAG  
2651 TAGGTGATGC CAGCCCTCAA GTCTGTCTTC AGCCAGGGAC TTGAGAAGTT  
2701 ATATTGGGCA GTGGCTCCAA TCTGTGGACC AGTATTTTCA CTTTCCCTGA  
2751 AGATCAGGCA GGGTGCCATT CATGTCTTT CTCTCCTAGC CCCCTCAGGA  
2801 AAGAAGGACT ATATTGTAC TGTACCCTAG GGGTTCCTGA AGGGAAAAACA  
2851 TGAATCAGG ATTCTATAGA CTGATAGGCC CTATCCACA GGGCCATGAC  
2901 TGGGAAAAGG TATGGGAGCA GAAGGAGAAT TGGGATTTA GGGTGCAGCT  
2951 ACGCTACCCC TAAACTTTTG GTGGCCTGGG GCATGTCTTG AGGCCAGAC  
3001 TGTTAAGCAG GCTCTGCTGG CCTGTTTACT CGTCACCACC TCTGCACCTG  
3051 CTGTCTTGAG ACTCCATCCA GCCCAGGCA CGCCACCTGC TCCTGAGCCT  
3101 CCACTATCTC CCTGTGACGG GTGAACCTCG TGTACTGTGT CTCGGGTCCA

FIGURE 1, page 1 of 2

```

3151 TATATGAATT GTGAGCAGGG TTCATCTATT TTAAACACAG ATGTTTACAA
3201 AATAAAGATT ATTTCAAACC ACCAAAAAAA AAAAAAAAAA AAAAAAAAAA
3251 AAAAAAA (SEQ ID NO:1)

```

#### FEATURES:

```

5'UTR:      1-44
Start Codon: 45
Stop Codon:  681
3'UTR:      684

```

#### Homologous proteins:

##### Top 10 BLAST Hits

	Score	E
CRA 335001101587561 /WO200058473 /org=Homo sapiens /taxon=9...	428	e-119
CRA 18000004937398 /altid=gi 464561 /def=sp P35289 RB15_RAT RAS...	423	e-117
CRA 18000005187045 /altid=gi 7498104 /def=pir T33855 hypotheti...	220	6e-56
CRA 18000004929618 /altid=gi 131798 /def=sp P24407 RAB8_HUMAN R...	216	1e-54
CRA 335001098683352 /altid=gi 11422744 /def=ref XP_001482.1  TR...	214	4e-54
CRA 18000005096141 /altid=gi 2317272 /def=dbj BAA21744.1  (AB00...	214	4e-54
CRA 18000004952869 /altid=gi 131848 /def=sp P22128 RAB8_DISOM R...	212	3e-53
CRA 335001098688905 /altid=gi 11432830 /def=ref XP_007682.1  RA...	211	4e-53
CRA 18000004945380 /altid=gi 131847 /def=sp P22127 RAO1_DISOM R...	211	5e-53
CRA 18000005163099 /altid=gi 7705849 /def=ref NP_057215.1  ras-...	210	6e-53

##### BLAST dbEST hits:

	Score	E
gi 12333507 /dataset=dbest /taxon=96...	626	e-177
gi 12120217 /dataset=dbest /taxon=96...	377	e-102

#### EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

##### From BLAST dbEST hits:

```

gi|12333507 brain
gi|12120217 epid_tumor

```

##### From tissue screening panels:

Fetal whole brain

```

1 MAKQYDVLFR LLLIGDSGVG KTCLLCRFTD NEFHSSHIST IGVDKMKMTI
51 EVDGIKVRIQ IWDTAGQERY QTITKQYYRR AQGIFLVYDI SSERSYQHIM
101 KQVSDVDEYA PEGVQKILIG NKADEEQKRQ VGREQQQLA KEYGMDFYET
151 SACTNLNIKE SFTRLTELVL QAHKELEGL RMRASNELAL AELEEEEGKP
201 EGPANSSKTC WC (SEQ ID NO:2)

```

#### FEATURES:

##### Functional domains and key regions:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

205-208 NSSK

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 2

1 92-94 SER  
2 206-208 SSK

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 2

1 29-32 TDNE  
2 104-107 SDVD

[4] PDOC00007 PS00007 TYR\_PHOSPHO\_SITE  
Tyrosine kinase phosphorylation site

101-109 KQVSDVDEY

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 2

1 18-23 GVGKTC  
2 136-141 GQQLAK

[6] PDOC00017 PS00017 ATP\_GTP\_A  
ATP/GTP-binding site motif A (P-loop)

15-22 GDSGVGKT

[7] PDOC00579 PS00675 SIGMA54\_INTERACT\_1  
Sigma-54 interaction domain ATP-binding region A signature

11-24 LLLIGDSGVGKTCL

**BLAST Alignment to Top Hit:**

>CRA|18000004937398 /altid=gi|464561 /def=sp|P35289|RB15\_RAT  
RAS-RELATED PROTEIN RAB-15 /dataset=nraa /length=212  
Length = 212

Score = 423 bits (1077), Expect = e-117  
Identities = 207/212 (97%), Positives = 209/212 (97%)  
Frame = +3

Query: 45 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 224  
MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ  
Sbjct: 1 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 60

Query: 225 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 404  
IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG  
Sbjct: 61 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 120

Query: 405 NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHRKELEGL 584  
NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHRKEL+GL  
Sbjct: 121 NKADEEQKRQVGREQGQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHRKELDGL 180

Query: 585 RMRASNELALAELEEEEGKPEGPANSSKTCWC 680  
R ASNELALAELEE+EGK EGPANSSKTCWC  
Sbjct: 181 RTCASNELALAELEEDEGKTEGPANSSKTCWC 212 (SEQ ID NO:4)

>CRA|335001101587561 /dataset=GENESEQ /org=Homo sapiens /taxon=9606  
/mol\_type=protein /date=08-FEB-01 /length=218  
/altid=derwent\_id|B41604 /altid=derwent\_ac|B41604  
/def=Human ORFX ORF1368 polypeptide sequence SEQ ID  
NO:2736 /patent=WO200058473-A2 /pat\_section=Claim  
Length = 218

Score = 428 bits (1088), Expect = e-119  
Identities = 212/218 (97%), Positives = 212/218 (97%), Gaps = 6/218 (2%)  
Frame = +3

Query: 45 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 224  
MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ  
Sbjct: 1 MAKQYDVLFRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIEVDGIKVRIQ 60

Query: 225 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 404  
IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG  
Sbjct: 61 IWDTAGQERYQTITKQYYRRAQGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 120

Query: 405 NKADEEQKRQVGREQGQ-----QLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHR 566  
NKADEEQKRQVGREQGQ QLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHR  
Sbjct: 121 NKADEEQKRQVGREQGQKCPSLQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQHR 180

Query: 567 KELEGLRMRASNELALAELEEEEGKPEGPANSSKTCWC 680  
KELEGLRMRASNELALAELEEEEGKPEGPANSSKTCWC  
Sbjct: 181 KELEGLRMRASNELALAELEEEEGKPEGPANSSKTCWC 218 (SEQ ID NO:5)

**Hmmer search results (Pfam):**

Model	Description	Score	E-value	N
PF00071	Ras family	323.8	8.2e-95	1
CE00060	CE00060 rab_ras_like	211.0	1.8e-59	1
PF00006	ATP synthase alpha/beta family	4.2	2.1	1

**Parsed for domains:**

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00006	1/1	10	24 ..	203	217 ..	4.2	2.1
CE00060	1/1	2	165 ..	16	184 ..	211.0	1.8e-59
PF00071	1/1	10	212 .]	1	198 []	323.8	8.2e-95

1 GCTCAAGATT GCACAGCTGG TGAGTGGTGA CACTGGGACT GGAACCCAAG  
51 TGTGCCCTTAC TCCAGAGCCC TTGGCATGCA CCTGAAACCC CATGTAAGCC  
101 CACTGTGGAG ACGCGCACCT CGAAATAATG GAATCCACTA CATCAGTTCC  
151 TTTAGCTTTC TGTGTAATCA GAGTAGCTAG CAGGCTCGGG ATTTTCGCCCC  
201 CCGGCTTTTT TTTTTTTTTT TTTTGTAGAC AGAGTTTTGC TCTTGTGTCC  
251 CAGGCTGGAG TGCAATGGCG CAATCTCGGC TCACCGCAAC CTTTCGCCTCT  
301 CAGGTTCAAG CAATTCTCCT GCCTCAGCCT CCCGAGTAGC TGGGATTACA  
351 GGCACCGGCC ACCACGCCCA GCTAATTTTT TTATATTTTT AGTAGAGATG  
401 GGGTTTCACC ATGTTGGCCA GGCTGGTCTT GAACTTTTCC CCTCTTATTA  
451 TAATTCAGAC ACTTAACCTG AAATATACCT TTTCAAATGA AGTAAATGGG  
501 CTTACCACTT TCCTTGACCT ACTATTGAAA AATACATTCT CCATCCAATA  
551 TTCAGCCTGA AAACAGGTAT GTACATATAT ACTTTTCATT GCTTTTTTTTT  
601 TTTTTTTTTT GAGACAAGGT CTCCTCTGTG TGCGCAGGCT GGAGTGCAGT  
651 GTCATGATCT CGGCTCACTG CAGCCTTCCC CTAATGGGTT CAAGCAATCC  
701 TCCCACCTCA GCCTCTCAAG CCTGGGATTA CAGGCGAGCC ACCGTGCCCCA  
751 GCTAATTTTT TTTTATTTTT AGTAGAGACT GGGTTTCACT ACATTGGCCA  
801 GGCTGGTCTC CAGCTCCTGA CCTCAAAGTG ATCTGCCCCG CTCAGCCTCC  
851 CAAAGTACTG GGATTACAGG CATGAGCCAA CGCGCCTAGC CTTTCATTGC  
901 TTTTAAAGA CCTAATAGGC TAGACTTTGC TCTCCCTCAA TACTCGTTGG  
951 TAGGGATAGG CAATTTTCTC AACTCCGGAG AGCATTCAAT TGCCTCTCTC  
1001 CGGTGCTAAC ACATTCAGTG GTAGGAACT GGATCTTGAA CAAGGGCCAT  
1051 TCATTCTTTG GTGCCACTGG CTATACCACA GAGAAATTTA GGGGTCTGAA  
1101 ACAATACATT GGTACCTGG GCACCTATCC TAAGCACCTT AGAGGGAAAA  
1151 CGGAGACTTG CCCGCACACC TCTAAAGGAT TTTGCACTTG GAGATGTTCT  
1201 TATGGCCATC TATCTTTTCA CCCTGGTGGA GGCCGTGAAT AGGCATTTTC  
1251 CCCATTTAAA GAAAAAATGG GGACGGGGGA GGGCCGTGAC ACAGTCACAC  
1301 AGGTAAGGGG CAGCCAGATG GCAGGGAGGG GGAATTCCAC CCACACTCTC  
1351 GGGGACTCAT GGAGACGGGT GTTCGAATCC AGATCCTGCT CAAGGCCTTC  
1401 CTACCTCGGG TGAGCCCAGC TGAGGTACCA GCCACTGGGG AGCCCGGCCA  
1451 GATCCTGCAG ATGCAGGGTG CCACGGCGGG CGGAATTACC GGCGCCAGAC  
1501 TTGGGGTGGG ATATGGGGAG AAGTGGTGAG CCCGGAAGC GGAGCACGGT  
1551 AGAAGTGGGC TGGGTGGGGG CTCACCTCAA CTCCCCATT CGGAGCGTCC  
1601 GCGGAAAAAC GAAACGTTT CCCC GCCCGG GGCAGGAAGG GGTGGGAGG  
1651 GGGGGCTGGC GCCCCGCCCC AGCGTCGCCT GCTCGATGGG GTCCCGCTCT  
1701 CCTGCGCGCG CTCCCCGCC CCTCTCTACC GGGCGCGCGG CGGCGCGCA  
1751 GGGGAAGGGG CGGGCAGGGG CCGCGCCCGG TTTCTCTCC CACCGCCTCG  
1801 CGCCAGCCCA GCCGAGCCGA GCCGAGCCGA GCGGGCGCGG CGCCGGGCTC  
1851 CCGCCGAGC CGCGCTTCCC GGCACCCAGC GAGCGAGTGG GCAGGCGGGC  
1901 GGGCGAGGCA GCGCGGGGG CCGGGCCCGG CGTCTCTCTC GCGCCCGCA  
1951 GCGTCCCCGG GCGGGCGCGG GCGCGATGG CAGCGCGGGA GCAGGGCTGA  
2001 GCCCGCTGCC CGCCCGCAGT TCCCGGCCCC GCTGGCCCCA GTCATGGCGA  
2051 AGCAGTACGA TGTGCTGTTT CCGCTGCTGC TGATCGGGGA CTCCGGGGTG  
2101 GGCAAGACCT GCCTGCTGTG CCGCTTACC GACAACGAGT TCCACTCCTC  
2151 GCACATCTCC ACCATCGGTA AGGGGCGGTG GCGCGGGCG CCCCTCCCTC  
2201 CCGCCCGCGG GCCCTTTTCC CCGCCGCCCC CGTCCCCAGC TGGGGAGGAA  
2251 TTGCCAGCCC CTCGGCTGG AGGCGGTGGC GCGGAGGCC GGAGTCCGGG  
2301 ATAAATCTCG GGGTGAGCAT AGGTTTGGC AGGTGAGGT GTCCCTGCTG  
2351 CCTGCCGTCC GGACAGGGG TGGGGTCTCC CGCCTCTTGC CGGGAAGCCT  
2401 TCCGTCCCAT CAAACCGAGA AACCAGGGGT GAGGGGAGCT GGTGTAGGCC  
2451 TGGGTACCCC GAGCTGGGT AGCAAGAATC GTAGCCGCTG GAATAACACC  
2501 CCCACACCCC CAGGGGAGG GGAAGTAAAG CTTCTGCTAC GGAAAAGGGG  
2551 GTCAGGGTGG AGACCGAGT CACTGAGGCG CCCTTGGTTC TGTGGTGACC  
2601 CAAGGTGGAG CCGGCGGGGG GCGAGGGGG GAAGAGAGGA CGTACGGAGG  
2651 GGCCACAGGG ATCGAGTTTC CAGGGCAGAG TTGGGAAGGT AAGCCGCAAG  
2701 GTGGGACACC TGGGGAGGA CACAGATAGG GTGAGGAGCC CCTGCGCCTG  
2751 GGAAGAGGAG ACATCTGTTC TGAGGGAGGC TAAAGAGGAT GGAGGAGTGT  
2801 CAGGAATACC TGCCAGACC AAGGGGTCAG AAGGCAGGCA GGACCCGCCT  
2851 GAGGGCATCT CTCATCTGGC AGTGCTGGAG CCTGTGGTTA GAGGGACAAG  
2901 ACCCGGTGGC ATCCAGACA GCACTATGAT GGGGTCACTT ATTCTAGGAA  
2951 TGGGTCCATG GCCTCCCCTC TGAGACAGTC AGTCTCCCGC TTCTAGGCTG  
3001 TGAGGGGGCC CCTCCCTGAG AAGTCTGAGT AGAGGGAATT TCATCCTCAG  
3051 CTGCTACCCG GGTCAGCCCT GGAGTAGCCT CTGCATTGCC CAAGCCCCTG  
3101 GAAACACCTG CTGGCTGGCT GGTATCCAT TTGGAATGCT CTCCTAGAAG

FIGURE 3, page 1 of 14

3151	TCCCTGCTGC	CATCAGGGAT	GGGCACCAGC	TCTCAGCTTC	CTCTTGAGGA
3201	TTCATGTCCA	CACCATCCCC	CCTCCCCCCA	ACACACATTC	CTTGCTGAGA
3251	GAGAAGTAGG	AGCAGATAGA	TACAGCCAGG	AGGAACAGAA	CCTTCTGGTT
3301	AAGAAGCCAG	CTTTATTGTC	CAAGAGACCT	GAGACCTCAC	TGTGGGGCAA
3351	AGCAACCTTG	AATATTGCCT	AACTTCTGA	GCTTTATTTA	GTTTCTCATC
3401	TGTAGAACGG	GTATAATAAT	TGCACCTACC	TGCCAAGTTG	TTGTCAAGAT
3451	TAAATGAGAT	AACGATTGTT	AAGTGCTTAG	CACAGCCAGA	CACATGGTGA
3501	AGCTCGATAA	ATGCTGATTG	TTCTTACTGC	TATTGCCATT	ATCATTGAGC
3551	TTTTAGGGTC	TCCTCTCTTT	GTTTCACCAA	CTTGAAGGGT	GAAACAACAG
3601	GACTTAGGGT	CAGGGAACAG	AACTTGTCCG	TCTTCTCAG	AGGAGCTGTA
3651	AGGCCAACTC	TTAGGAAACC	CAGGAGCTTG	GGCTGAGCCA	TGGTTTGGAT
3701	GAGAGACATT	GCAGAAAGAA	GGGAGCCTA	TAGACACTAA	GGCTTTGTGC
3751	CTGCCGGGAG	GACTTGGGGA	AGAGGCAGGT	GCAGGAGAAA	GGCATGGGCG
3801	TGATGGAGGA	AGTGGCAGAG	GAACCAGATG	GTGTATGAGG	ACAGGTTGTG
3851	GGCTCAGGGA	CAAAGGGCGG	TGGGTATACC	CTTAAGGAAA	CTAGGAGTGG
3901	CTATTTTTTG	GAGAGGCCTG	GTGCTTGGA	CTACTGAGCT	ATCTCCAGAG
3951	AGCTGTGGGC	TGCCTGGGAG	GCCCTGGCTT	TGGCCTGAAG	AGCTGTTGTT
4001	TGCACCTGCT	CTCCTAGTCC	CATTCCAAGT	CCTATAGGTG	ACATGGACTT
4051	TTCCCTTTGA	GGGCTTCATT	CAACCACCTC	ATTTCAGAAG	CTCTGGGACT
4101	CCTGCTTAGT	GCTGTGGGAG	GCAGCCTCCC	CTGGGAGACA	CATACCCTCC
4151	TTTTTGAGGG	CACCCCTCTT	TCTAAAATAC	CAGGATGGCC	CTCTGAGGCT
4201	CGTGCTCTCC	TCAAAGAGAG	TCCATTGCCT	CACACCTCTA	ATCATCCACC
4251	CTTCTCCTTG	TCCCTTCCCC	TTGTAATCTC	CCTTCTTAGA	CACCTTCTGC
4301	TAATAGGTGA	ACACTAAATA	GGTCACAGGG	ACTTCCTGAA	ACCCTCCAGG
4351	GCAGACCAC	TTGGGCACAT	AGGTGAATCA	GTGAACTGAG	TAGGGGTGTC
4401	TCTGCAGCAC	TGTCTCCCTT	CAAGGCCCTT	GGTATATTGG	CCTAAAACCT
4451	AAAGATGGCT	CCCAGATTTC	TTCTTCCGCT	TCTGACACCC	GGGTTCCCTT
4501	TTCTACAGGA	CACAGAGGAT	TCTCTAGGGT	CCCCCTTTCC	ACAGGACACA
4551	GAGGACTCTA	GAGAGTTTGA	TTCCATGGAA	TAGAAAAGAA	CCTGTCTTTC
4601	TTACACACCAG	CCTTTTAAAA	TCTGCCCCAC	TGGGTATCTT	AAATGCTTTC
4651	TTATTTAAAG	CTTATTAAGG	GACTTGGGAT	TCTCCCTTAT	CTTGGGCGTG
4701	TTTTTTCAGCA	TTAACTAAAA	CTTAAAGGAA	AGAGTTGGAT	GGTCAAGAAA
4751	AGCTTTTTTC	TTAAGTGATA	TGGACAGTTT	CTCAAGGAGG	TAGAAGGGGC
4801	AGCCAGGAGA	GAAATCAAGG	AGCCAACGAA	ATGAGTGCTA	CCAAGTCATA
4851	GTCAATCGCT	TATTTTTTAAA	AAATGCGTGT	CCTGTATGCC	AGGCTCTGCA
4901	CTGAGACCGA	GAGATTCCAA	GATGAATAAT	ACCTACAGTC	ACTGTTCTCA
4951	AATTTGTGCAT	TACCTAAAC	ACATTACATG	ACCATGCTGG	CCACTGATCG
5001	AGGCACCTTT	CCCAGGGGCT	TTTTTTTGTA	ATTAAGAAAA	CAAGGTAATT
5051	CACCAGTTAT	TGCCAAGATA	GTTTGGCTTC	TTGGCTCATG	TGGATATCAC
5101	CTAGGCCAGT	ACTTTTGTA	TTTACTGTGT	ACTCCACTTT	AACGGCCTGC
5151	GATCTTCTAG	AGAGAAGCCC	GCCAGGGAGC	AGTGAGAGGC	CTCCCTGGTA
5201	GACTGAGACA	CTGACTGTCC	CTCCCCCTAT	CCTTTTCTGC	TTTCTGGCCA
5251	GCAGACCAGC	AGGTGGCCCT	GCCACTGGCT	CTGCCACAGG	CATTTCTTTT
5301	CTGTGCAGCT	GTGCTGGCCT	GGCTGGGGGT	TGGTGCGAAG	GGGTCCCCAA
5351	AATACTACCT	TAAACAAATT	AATTGAGCAT	TCACTACCAA	GCTCTGTGCC
5401	AGGCATTTTA	GAGACATATT	GCAGTCTACG	TTTCTGCCC	ACAGAAGCCC
5451	ATAACCTAGA	TGGGGAGGCA	AGACAAAGGG	AAAAACAAAA	AACAAAGAGC
5501	TAGTGCCAAA	ATGAGATATC	TGAAAGAACT	TGGTGAATCA	CTCTTCAAAT
5551	GTAAAGGATG	GATTATGATC	ATTGCAGTTA	CTCTTAATGA	AGGTCTCACA
5601	GTGGGTATCA	GAAGCTAAAT	TATGATGCAA	GATGTACCAT	GAGGCAGCCG
5651	GAGAATGGCG	ATGGATGGGA	TGGGTGAGTG	CTATTCCCAC	GACTCCATGC
5701	TGTCGGAGGC	TGGGGAAGAG	AGAGGCCCTT	GTGGACTAGA	ACCGGCAGGG
5751	AAGGCTGAAG	CTAGGCCTCA	GTGTGGGCTG	CTCGTCAGTT	CCTGCAGCAG
5801	AAGGGAGCAG	ATGGAGTAAC	ATGAGCAGAG	ATAACAGAGG	TGGGATTGAG
5851	TAGGTGTCCG	TGGGGCTCTA	GGCAGTTTAG	ATGCAACAGA	AGGGATTCTT
5901	CAGGAAAGTG	AGAAGATTCT	TCTGTTTCTC	TCTCTGTCTC	CCAAATTATA
5951	AGTGCCCTGA	TGGTGCGACC	AAATCTTATT	CCTCATTTGT	TTTATAGTCC
6001	CTAGTACAGG	GCCAGGCAGA	TTCAATGCCT	GTTGTTAAAT	TAATGAATGA
6051	ATGCAGGGAC	CAGTTGGCAG	AGGGCATTGA	GAGCCTGGCC	AAGGAGGTGG
6101	AACATGAGCC	TTAGCAATGG	TAGGAGGGGT	TTTGAGTAGG	GTACTAATGA
6151	GGTTGGCTGG	AAGAAGGGGT	TAAGACTTGA	AGCAGGGAGA	CTAGTCAGGG
6201	GCTGCAGTAG	TATCCTGGGC	ATGAAGGAAC	CTCTGAATGG	CCCCTCACC
6251	CCAGTGGTAC	CAACACCAAC	TTCCACACAG	TCAGTTGTTT	TACTTTCCCT

FIGURE 3, page 2 of 14

6301 CCAGAATGGG GAGTGGTTCA AGCCAATCAA CCTGGCAACT TCTGAAAGAA  
6351 TCTTATGGGA CCTGTGCCAT GACCAGGTAG GGAGAAGATG TCATACATGG  
6401 ACATCTATGT TCAGGGGACC TTTGAGGACC TTTCTGCATG GTGGCCAGGA  
6451 GTGAGATGAT GTAAACCACA AATGGAAACT GAAGAGACTG CTCAGGAGTT  
6501 GTTGGTTTTT TTTTCTTTTC TATTTTTTTT TTTTGTAGAC TAGGTTTCAC  
6551 TCTGTCAACC AGTCTGGAGT GTGGTGGTGG CACAATCACG GCTCACTGCA  
6601 GCCTCGATCT CCTAAACGCA ATCCTCCCAC CTCAGCCTCT CAAGTAGCTG  
6651 GGACTACAGG TGCATGCCAC CACATTCAGC TAATGTTTGT ACGTTTTGTA  
6701 GAGATGGGGT TTCACTATGT TGACCAGGCT GGTCTCGAAC TCCTGGACTC  
6751 GTGATCCACC AGCCTCAGCC TTCCAAAATG CTGGGATTAT AGGCGTGAGC  
6801 TACCTCACTC CCTCAGGAGT TGGTTTTCTC CCTCCCATCC TTAGTCTTCC  
6851 CTGAGTAGAC CTGTACCTA GTCCCTGGAC CTTTTGTTTT GAAAGCCACC  
6901 CTCCAGGCTA CACTCCTTCT GGGTGAGGAG GAGGGTGATC TGGTTGGACA  
6951 GGTGGGGCTG CTGTGGCTTC AGGGCACTTT CTCAGGCTGG GTTGCTGCTG  
7001 CTATGTCACC TTTCTCAAGG AGTTCTGCTG GGACTGGCTT GGCTGCCTGT  
7051 CTTGACTTTG CTTTTGACTG AGGAGGTGGG AGATGGTGAG GGAGGGGGTG  
7101 GGGCTAGATC CAAGCCTGGA ATGGGGTGAC CTAACAGACA CTGGGGCCTG  
7151 TGCTTAGACA CTAGGATCCT GGGGTTTGCA GGTTCCTAGA CTGAGAGGAG  
7201 CTGGGGGCAA ATGCAGTGTG ACGTTGTGAG AGGGTCAGGG CTGGGTCTGT  
7251 GTCAGCCTTC AGGCAGCCTG AGACCAGTCT CTACCTACTC TGTTCCCCTG  
7301 GTACCTAGAA AGGAAGGGAA GGTGAGAAGC AATGAGCAGA ATGGAAAGAG  
7351 CCCAGATTAA CATGCACATT TCCCATGGCC TTAAGTGGCC TGTGACCTTC  
7401 AGACACTTTG ATGACATCTT TGTGCTTCGT TTCTGCATCT GTAAATTGAA  
7451 GATGGTAACA GAGTCTTTCT TAAAGGTTGT TGTGAAGATT ATAGAGCCTA  
7501 GCGCATATAA AGCACTTGGC AGAGCCCTCG ATAAAATAAT AGCTGCTATC  
7551 ATATTATCAT TATTATTATT TTATTATTAT ATTTATTAT TTTTTTTTGA  
7601 GACCGAGTAT CTCTCTGTCTG CCCAGGCTGG AGTGCAGTGG CACAATCTCG  
7651 GCTCACTGCA ACCTCCATCT CCCGGGTTTA AGTGATTCTC CTGCCCTCAGC  
7701 CTCCTAAGTA GCTGGGATTA CAGGCACCCA CCACCACACC CGGCTATTAT  
7751 TATTATTCCT AGCTATAAGA ATGCTGTAGA GATGAATACA CTGTCAGTGA  
7801 GCTAGGAGGT CATCCTGTGT ATCCATCACT TGTGCACTCA GTCGTTCAAG  
7851 CGCTATTTGC TGAACACCAA CTACATGCCA GGTGCCATGC TAAGATTTGG  
7901 GGACACAGTG GTGACCAAAA CAGACAGAAA CCAAGGAGCT GGCTTACATT  
7951 CCAAGGGAGT GCATAGGAAG CTGTGTTTCA TTTCAAGTTT TGCTCTAGTA  
8001 CCCCCTTTT CCTGGCAGTG CCAGGGTCTG AGAAGGAAGA GTGAGGTGGT  
8051 GAGGAGGTGT GAAGCAGTGG GGTGACCTGA GAGGAGAGGA TGGGGTGGCT  
8101 TTGCCTCAAG GCTTGGGCCC CTGCTAGGTG TCGCTCTGCC TCAGGCCTCT  
8151 GTTTCTCCTC CTGACACAGG CACAGACTCG GCCTCCCACC CCTTCCCCAA  
8201 GGACATGACC TTGGGAAGGA ACATATCTGA AGCCCGCGGA GGGTTTCCGC  
8251 TGCTGTGCAT CTGTGCCACA GATCCGCAGA TGCACCCACA GCTGGGAGCA  
8301 CCGGTTCCCT CCGCCTACCT GCACCTCCTG GTTTCTGTTC CTTCCTCCTC  
8351 CTCCTTCCTT CTCCCCGCTC CCCAGACAGG CTGGTGATGA GCTTTATAAC  
8401 ATGAAAGCTG ATATTTGGCC ATTATCCTTC TACCTTGATT GCCAGCTCTT  
8451 CTCAGAGTGC CTTCTTCTGT AATCCAATCT TTGCACCAGT TTCCCTGTGA  
8501 AACTGCCAGT TTTCTGTATA GGCCTCTGCC CTCTCCTTGG CTCTTCTCTC  
8551 TGGTCAGTGA GCTTTGTCAA GGGGAACACA GGGCTTCCTG GACACGTAAT  
8601 TCCTCCCAC T GAGGAGGAAG GGGCTAATCA CCAGCCCTGT TTTATTTTAT  
8651 TTTATTTTTT TGAGATGAAG TCTAGCTCTG TCGCCAGGC TGGAGTGCAA  
8701 ATGGCTCGAT CTCGGCTCAC TGCAACTTCT GTCTCCCGGG TTCAAGCGAT  
8751 TCTTCTGCCT CAGCCTCCTG AGTAGCTGGG GATTACAAGC ATGCACCACC  
8801 ACACCTGGCT AATTTTTTGT GTTTTTTAGTA GAGATGGGGT TTCACCATGT  
8851 TGGCCAGGCT GGTCTCGAAC TTCTGACCTC AGCTGATCCA CCCACCTCGG  
8901 CCTCCCAAAG TGCTGGGATT ACAGGAGTGA GCCACCATGG CTGGCCGACC  
8951 CCATCTCTTA AAAAAACAAA AAGAAAAGAA AAGAAAACAA AACAAAAACA  
9001 CTTTTTAAAT TAACGATTA TGGTGGCATG TGCCTGTAGT CCTAACTACT  
9051 CAGGAGGCTG AAGTGAAGG ATTGCTTGAG CCAAGTAGT TGGAGGCCAC  
9101 AGTGAGCTGT GATCACACCA CTGTACTCCA GCCTGGGTGA CAGAGTGAGA  
9151 CCCTGTCTCA GGAATAAAAA AAAATTACTG AGAACTCTGT GACCATGGCA  
9201 CCATGAAC TAAGAAAGGGC TAACAGTTGG CTTTGAAATG TGGGTTATGG  
9251 CTGGGTGCGG TGGCTCACGC CTGTAATCCC AGCACTTTGG GAGGCCAAGG  
9301 TGGGCAGATC ACAAGGTCAG GAGTTTGAGA CCAGCCCGGC CAACATAGTG  
9351 AAACCTCATC TCTACTAAAA ATACAAAAAA TTAGCCGGGT GTTGTTGGCAG  
9401 GTGCCTGTAA TCCTAGCTAC TCGGGAGGCT GAGGCAGGAG AATTGCTTGA

FIGURE 3, page 3 of 14



9451 ACCCAGGAGG TGGAGGTTGC CACAAGCTGA GATCGCACCA CTGCACTCCA  
9501 GCCTGGGCGA CAGAGCAAGA CTCCATCTCA AAAACAAAAA TAAAAACAAA  
9551 AAAAAGTGGT TTGTTTTCTT TTCTTTCTTT TTTCTTTTTT TTTTTTTTTT  
9601 TTTTGAAACA GAGTCTTGCT CTGTACCAG GCTGGATTGC AGTGGAGGAT  
9651 CTCAGCACAC TGCCACCTCT GCCTCCCAGG TTCAAGTGAT TTCCCTGCCT  
9701 CAGCCTCCAG AGTAGCTGGG ACTACAGGCA CGCACCACCA CGCTGGGCTA  
9751 AGTTTTTGTA TTTTAGTACA GAAGGGGTTT CACCATGTTG GCCAGGATGG  
9801 TCTCCATCTC CCTGACCTCG TGATCCGCCC ACCTCGGCCT CCCAAAGTGC  
9851 TGGGATTACG GGCATGAGCC ACCACGCCCC GCCTAAAAGT GGGTTATTTT  
9901 CTAATTGCTC TTCCCTGATT AAAATTTTCT CTTTGCCCAT CTTTTCTCTA  
9951 GATATGTACT GACTTCATTC ATCCATTTAT TCGTCTCACT TGCTCATTCa  
10001 TTTTTGCTTT CATTACTCT ACTTTGTTGA ATAATATTTA GTGATCTACC  
10051 TGCTGCCAGG CAGTGAGAGT CTGAAGTGAA CAGGATGCTG CTTTGCCCTC  
10101 TGGGAGCTTA CAGTGTAGCT GGAACACAGA CATCCAAACA AGCAGAATAT  
10151 TATGCAAAAG AATGTACAGG ATGCTTTGGA ATCACAGAGG AGTGAGAAAT  
10201 CCCTCCCCGG GAGGCTGGTG AAGGCTTTGA AGAGGAAGTG ACATTTGAGT  
10251 GGAGTCTTGA AGACTAGGCA GGATTCTCCA GGGGCCCTGG GTGTGGGGGA  
10301 AGCACACATC CTCTTCCCTG TAGGAGGTGC TGTGGAGAAC ACCTCCAGTG  
10351 GGGCTGCTAC TCTTCAGCCT TGCTGGGGCC AGCTGGAGTG GCCACACCAT  
10401 GGTCACACCA GCTGAAGTTC AAGAAGCCCC TTGCCAGGAG ATTGCTTTGC  
10451 TGGCTCTGGG TGAGGGCAGG TGCATCTGGA AGCCCCCTTC TTTCTAAGAT  
10501 GTTTGCTCCT GAGTTTCTAT GTCTAGTCT TTTCTTCCCT GAACCTTTTG  
10551 CTACCAGTCA GCACAGCCCT GCCTGAGAAG GAGGCTGGAG GAGTGAGTGG  
10601 TCAGTAGCCT GGTGGGTCTT GGCTGCCTCT GTGGTGCCCG CTGGCCTAAG  
10651 TAGCAGGCTT AGGGAGGCGA GACCCAGTTC CAGGGGCTGC CAATGGGGAG  
10701 CGAGATGGGG TGGCTGGAGC ACACTGCACA TGTCACCAAG GCTCTAGGGA  
10751 GGTCTGTGCA CAAGGCAGTG GGAAAAGCAA GGGGAAGACC CAGCCTGGTC  
10801 AACATGGTGA AACCCCGTCT CTAATAAAAA TACAAAAATT AGCTGGGTGT  
10851 GGTAGAGCAC GCCTGTAGTC CCAGCTAACT TGGGAGCCTG AGGCAGGAGA  
10901 ATCACTTTAA CACAGGAGGT GGAGGTTGCA GTGAGCCGAG ATCGTACCAC  
10951 TGTACTCCAG CCTGGGTGAC AGAGTGAGAC CCTGTCTCAA AAAAAAAAAA  
11001 AAAAAAAAAA AAAAAGTGGG GAAGGGGAAC ACTGATCCTG ATTATCTACT  
11051 CCATATACTT ACTATGTACC TACTACCTAC ACAGGGACGG TGGGCTTTAC  
11101 GCATGCCATT TATTCACTGT ATAGAGATCT CAGCATCACA TAGGAGCAGG  
11151 GAGTTCTGAA GTTGGCCTTG CTGGCATTG AGAAGTTTCT TGGTGTATTC  
11201 TTCAGGTTCA CGCCTCCAGA CAAGTGTAAG TGCTATTGAA TGCTGACTAT  
11251 GTTCCAGGAA CTAACCAGA TGCTAGAAGA CACGCAGTAA ACAGTACAGA  
11301 TGCAGGTGCA CATGTGAGGG TCCACACAAG ACCTGAGAGA AGGGAGGGGT  
11351 CTTGCTGCAG TTCCCTTTTT GTAACAAAGG AGAGAGTACT GTTGACCTC  
11401 TTCCTAGGAA CTGTGAGTTT GACTGAAATG TGTCTGCCA CAGGATCTTT  
11451 GCTGCTTCTT CTACCTGATT CTTTGGATCT CCCTGCTGGC ACCTTCTTGT  
11501 CATTTAGGTC TCAGCTCAAA TGTTACCTCC TTTAAAATGT CTTCTCTGGC  
11551 CAGCCAGTCT AAGGTTGCTT GTGCTTGGGG TCTCCTCACT CTCTACTTTA  
11601 TCCCGCAGTT GCTTCTTATC ACATATGGCT CTCTGAAATT AGGTATTCTAT  
11651 TACTTACATC TGTCTTCCCC ACTAGAATTA AGCTCTGATG ACAAGGATCT  
11701 TTCTGTGCTG TTCATAGCTT ATCTTCTAGT ACCTGGCTTA GTTCTTGCCA  
11751 CATTGTAAGC ATTCAATAAC AGTTTGAATG AATGAATTAA CAAATGAAGG  
11801 AATGAATGAA TGCATTTTCC TAGAGGACTT CTGTTCTTCC CTGAGGGAAG  
11851 TTATAGGTCG TATTGGTTTC TTGGGACTGT TTTTGTGTTG TTTGTTTTGT  
11901 TTTGTTTTTT GAGACAGAGT CTCACTGTAT CCCCCAGGCT GGAGTGCAGT  
11951 GGCACAATCT TGGCTCACTG CAACTTCCGC CTCCCAGGTT CAAGCGATTC  
12001 TCATGCCTCA GCCTCCCAGG TAGCTGGGGA TTCCAGGAGC CTGCCACCAC  
12051 GACCAGCTAA TTTTGTGATT TTTAGTAGAG ACAAGGTTTC ACCATGTTGG  
12101 CCAGGCTGGT CTTGAACTCC TGACCTCAGG TGACCTGCCT GCCTCTGCCT  
12151 CCCAAAGTGC TGGGATTACA GGCATGAGCC ACCACGCCCC GCCTGTTTTT  
12201 TTTTTTTTTT TAAGACAGAG TCTTGCCTG TCTCCAGAC TGGAGTGCAG  
12251 TGGTGTGATC TCAGCTCATT GCAGCCTCAA CCTCCTGGCC TCAGGTCCAG  
12301 GTGATCCTCT TACCTCAGTC TTCTGAGTAA CTGGGCCAC TGGTATATAC  
12351 CACCACACCT GGCTAATTTT TAAATTTTTT GCAGAGACAT GGTCTCACTA  
12401 TGTGCCCCCT ACTGATCTTG AACTCCTTGG GTTCAAGTGA TCCTCACACC  
12451 TTGGCTTCCC AAAGTGCTGG GTTTACAGGT GTGAGCCACC ATGCCTGGGC  
12501 TTGAGACTGT TAAGATGATG AGGCTGGAGG GAGTGGATGG CCTCACTGCT  
12551 TGAGCCCTAG AGATTCCCTA CTCCAGAGTG CCCTGGCTGC AGAGGTGGCC

FIGURE 3, page 4 of 14

12601	CTGGAGGGTC	ACTCCAGCAA	CCTGGCTGAG	CTGATGGGCA	TCATCTGATA
12651	CCAGCTCTGA	CCCTGAATAA	TAGGCAACAT	GGACCTTAGT	CTAAAGCACT
12701	GACCCCTCAT	CTCTGCATAT	ACCAAAGAAG	ATGAGATTGT	GGTGAGGACA
12751	CAGCCAAACC	ATATCAGCTC	CCGGGATCCC	TGTGTGAATG	GGGTCTTTTT
12801	TGGTGTTTGA	GGGCTGCACA	GGGTGACCTC	TTTAGAGGTG	ACCTCCTGCC
12851	ACAACCCACA	GGAGGTGCAC	ATGGCCACAC	CATGCTGGTT	TCCTGCAGTG
12901	GGAGGGGCTG	GGGCACTCCT	GGGACCTGTG	CTTGGTAAC	GGAGCTGGCC
12951	TGGCCCTGGG	GATTGGGTGT	CTGCCCTGGG	TTTCAGGTGT	ATTAGGTTGT
13001	TCCTCGTTGT	GGAGTCTCAT	TACTAATGAA	AAGTTCAGGT	CGCACTGCTG
13051	GTCCTTTGGG	CTGTGGTTGA	TCCTGGTGAT	AACATTTGGC	ACCCAGAGGC
13101	AGCCCTGTTT	CCACTGAAGC	ATGCGGAGCT	TGGCTGGCAG	GCAGGCAAGC
13151	TGGCAGCTGC	CCTTAACCCA	TGAGGTGCTG	GCCCGCTAGT	AGGCACACCC
13201	TACCTGTGCC	AGAATTGAGG	TTGTAGCCAG	ACTCCAGGAG	CCATCTGGGC
13251	CCCACAGGGG	GCGGCATTTT	CTCTTTTGT	TGAAACATTC	CAGCCAAGTG
13301	CTGGCTTGGG	CTTCATCTCT	CTGTCCCCT	CTCCTTCCTC	TCCCCAACAT
13351	AAGCCTCCTT	CTACATCCTA	GAGCTCTTTC	CATTCCCCCT	CCTGCAGCTC
13401	TGGGCTCGCT	AATCTCATGC	TTCCCTAAGG	GAGCTTGACG	GCTGCTTCTG
13451	CTAACATTTA	ATAAAGTTCT	GCGTGCCAGA	CCCTGTGTTA	TGGGTTTAC
13501	ACCTTATCTC	ACAATCTTAA	AAAAAAAATT	CTCTGAGAAT	CCTCTGTCAC
13551	CCCCACTTTA	CAGGTGAGGA	AACCTGAGCA	AAGATAGGCT	AACTGGCTTC
13601	CCCAACACCA	TGCAGGTAAT	TAGTGATAAA	GGCAGGGTTG	GAACCAAACT
13651	TGACCTCCCA	ATTGTGCTCT	TAATGGCCAG	GACACTCTGT	GTCTTGAGCC
13701	ACACTTCCTC	CATGTTTTCT	AGGGCTTTCT	AGGGAGGCAG	ACAGTGATGG
13751	GAAGGGGTGT	TCTTTAGTGT	GGATGTGCC	TGCCTGCTCC	TTTCTGTAAG
13801	CGTCACAGCA	CCTCCACTGC	TGTACTGGG	AGGCACCAAG	TTTTTCCCTG
13851	TTTGCCCAACC	CAAGCGCAGC	TAGCTTAGGA	GTACAGTGAG	TGCTGGGTGT
13901	CTCGCTGTCT	GCATCCCTCT	ATCCTGCCCC	TGCCCCGGT	GCCCAGAGGA
13951	GGGCCCTGCC	TGTCTTCCCA	GTTCTCCAAC	AGCAGCGCTG	TCCCAGCACC
14001	CTCGGGCTCC	AGTTGTGGCC	TGGCAGCTGC	TGGGGCAGAC	ACCATACAGA
14051	CAGAGTCACA	GCAGGAAGAG	GATGGGGCCC	AGGGCTGCTG	CCTCAGGCCA
14101	TGGCTGCATG	GCACCATCAG	TTGATTGAGG	AGCTTTTCTT	GCCAATGTCT
14151	GAGGCATCAG	GTGGCAGGAC	ACGTCTCCCT	GCTCTTAAGC	CTCAGGCATG
14201	CAGCCCTTCT	TATGCTCTCT	GGGGTGAGGG	GGAGATCCCC	CTCATGGAAT
14251	TGCTTTTCTT	TTTTTTTTTT	TTTTTTTGG	ACAGGGTCCT	GCTCTGTCAC
14301	TCAGGCTGGA	GTGCAGCCTC	AACCTCCCAG	ACTCAAGTGA	TCCTCCTGCC
14351	TCAGCCTCCC	GAGTAGCTGG	GACCACAGGT	GGACACCATC	ACACCTGGGT
14401	TTTTTTGTGT	TTTGTTTTTT	GTTTTCTAGA	GATGGGGTCT	CACTTTCTTG
14451	CTCAGTCTGG	TCTCGAACTC	CTGGGCGCAA	GCAGTCTCTC	CACCTCGTCT
14501	TCCCAAAGTG	TTTGGATTAC	AGGTGTGAGC	CACTGTGCTT	GGCCTTTTTA
14551	TTTATTTAGA	ATTTGTTTTG	GAATTGCTTC	TTTATGCTCTG	GCACATAGCT
14601	GGCACTATGT	GGCAGAGATT	TTAAAAACGA	GCAAAACAAA	CAAATGCTTT
14651	GTCAACCACA	AAATGTATTC	TCTGCCCCCT	AGGTTCTTTT	TGTGTAGTTG
14701	AGGCTAGAAG	ACAAAAATAG	GGGGCAGTAA	GGAGCAGGGA	GCGATGGTTT
14751	AGGAGGTCTT	CCTTCCAGCC	CCCTTGTTGA	AGCATCTGGC	TCACTAGCTT
14801	GGGGGAGCCA	TTAGGCAGCA	GTGGCCAATC	CTGAGGCACT	CTCAGGTGTC
14851	ACTAAGAAAA	GGGGCATGTG	CTCTATGGAT	ACCCATGGGC	TGAACCTGGA
14901	GTCTGGTCTG	GACCCATGGC	TGTGCTAGGA	TCCACCGTCC	CCAGCCCCAA
14951	CTGCAGTCAG	CATGTTTCATC	ATCCTTAGGC	CTCTCCGCTT	CTTTCTGCAT
15001	GTTTGTCTGC	CTCATGCCCT	GCTCATTACC	AACTGGTCAG	TCCCCACTGC
15051	CCTGCCTGGA	GTGAGCTGGT	TTGATTGGCT	TAGCTAAGCT	CCCTTGCCCTC
15101	TGCTGGCCAG	GTCAACCCTGT	GGGTCAACAG	CAAACCTGTT	GATGGTCCAG
15151	TCTGAACCTG	CTTCTCCACA	AAGAACATGT	TGCACCCAGC	CCTGCTTCTC
15201	TGAGCAGAGG	TTTGGGGCTG	AGCTGTTCTA	GCCAGAAAGG	GACACAGGGT
15251	GTGGCAGGCA	CCATGATGGG	CATATCTAAT	GTGCCGGGAA	AAACAATGAG
15301	CTGCTCTCCG	TGCTTTGGGC	ACCTGGTTGG	GAGAGGGCCC	ATCTGTCTGA
15351	CTTTCTCCTC	CTGGGGCTCT	CAGCGTCTCC	GAGAACCTCT	GCCAGAGCTG
15401	TGTAGAAGTG	GTTTGCTTGT	TTCTTAACAC	TTCTGTGCCC	TATTTCTTTC
15451	TGTACCAAGA	AAAGGAAGTA	GACTGTTTGT	TAGGGACACT	GTGCGGGTGA
15501	TGAATCTGGA	CTTACTGGAA	TCATGAACCA	TGCCAAGGAG	GAAGGAGAAA
15551	ATAGGCTATG	GTGGGTGTCT	TAGTTAGGGC	TGGCTGCTGT	AACAAAATGC
15601	CTTTAGCTGA	GTAATTTAAA	GCAAGAGAAA	TGTATTGCTC	AGAGTTTGGG
15651	AGGCTGGGAA	GTCCAAGATC	AGGGTGCCAG	CAGATTAGT	GTCTGGTGAA
15701	GGCTGATGCT	CTGTGACAAA	GGTGGCACCT	TCTAGCTCCA	TCCTCACATG

FIGURE 3, page 5 of 14

15751 GCAGAAGAGG GAACAAGCTC CCTCAGACCT CTTTTCTAAG GCGTTAGTC  
15801 CCATGCATGA GGGCTCTAAC ATCACGACTG AGTCACCTCC CAAAGCCCTC  
15851 ACCTCCCACC AGCACTGCAC TGGGGATTAA GTTTCAATAT GGAATTTTG  
15901 GAGGAACACA GACCTTCAGA CCACAGCAGC GGGCTTCTCC TCATGTGCCC  
15951 CCTGCCTCAC TTCTAGATGC CGCATAATGT CAGTGAAACC CCGTCTCTAC  
16001 TAAAAATACA AAAAATTAGC TGGGTGTGGT GGCACGTGCC TGTAAATCCCA  
16051 GCTACTTGGG AGGCTGAGGC AGGAGAATCG CTTGAACCCA GGAGGCAGAG  
16101 GTTGCAGTGA CCTGAGATCG TGCCACTGCA CTCCAGCCTG GGCACAGAG  
16151 GAAGACTCCG TCAAGAAAAA AGAGAAAAGG CATCAGGTAT GCCAGGGTGT  
16201 GCGGGAAAAG GCATCGGGTA TGCCAGGGCG TGTGGGAAA GGCATCGGGT  
16251 ATGCCAGGGT GTGTGGGAAA AGGCATCGGG TATGCCAGGG CATGTGGGAA  
16301 AAGGTGGTAA GATTCTCTAG CCTCCAGGG TTGGGAAGCC TCTGGCCGAG  
16351 TGAAGCATAC CCTGGGTGGG TTTAAGACA CCAGCTTCC AGTCCAGCTC  
16401 AGCTGTGGGA TGTGGGAACA TGAGTCAGTG GGAACATGAG AATTGGCTTC  
16451 CCTGTGGCTC ACAATAATAC CTACTCTGCT CTACTTCATG GGACCCGCAT  
16501 AAGAGCTGAG GGATTCCATA GCTCAGGGGT ATGCTGTAAA GACAAGCACT  
16551 ATGCACCTGG GTGTGGTTCT GAAACTTTCT TGTGCAGAAG AGTGAGTAGG  
16601 GCTGGGCGAG TCCTGAGAAAT GTGCATTTCT CACACACCTC TGATGCTGCT  
16651 GATGCTCTAG TCCCTTGGCT GGCAAGGGTA CCTGGTTAGT AGGGGCCAGG  
16701 ACTCTGTAAT GCCTTCCACT TCAGGGTTCT CTGGGCTGGT TTCTCTGACT  
16751 CCCAGGAAG CTTTTATTCA GCAGAGGGA GGTAGGAGTG AGAGGACTAC  
16801 GCTGTCAAGT CTTACATAC ATCGTTAAT TTATCCCAGC ACAGCCCTTA  
16851 GGAGGGAAGC AGTATTCTCC TTCTACACTT AAGAAAATCG GCCTGGTGCG  
16901 GAGGCTCATG CCTATAATCC CAGCACTGTG GGAAGCTGAG GCGGGAGGAT  
16951 CGCTGGAGCC CAGGAGTTCA AGACTAGTCT AGGCAATACA GGGAGACCTC  
17001 ATCTCTACAA AAAAAAAAAA AATTAGCTGG GCATGGTGGT GCACACTTGC  
17051 AGTCCCAGCT ACCTACCCAG AGGCTGAGCT GGGAGGATTG CTTGAGTCCT  
17101 GGAGGATCGA GGCTGCAGTG AGCTATGATT GCTCCACTAC ACTCCATCCC  
17151 TGGCAACAGA GTGAGACTCC ATCCCAAAAA AAAAAAAAAA TTGAAGCTAG  
17201 GAGAAGTTGA GACTTGCTG AAGTTACACA GTAAGTGCCA GAACCAGGAC  
17251 TTGGACCAGG TCTTTCTGAC TCCAGGCCAA TGGATGTTT TTCCATGACA  
17301 TATATAGCTC TTGAAACTAC TTCTATCTAA TATCACCCAC AGTGCTGTTA  
17351 AAAATACAGA TTTCTGGGCC TCACCCTCAA ATTATGATT AGTAGGTCTA  
17401 GGCACGTCAA GTTCATTGTT TTTGTCTTTG TTTTAAGTCA CCCAGGTGA  
17451 TTCTAAAGCC GAAGCTCTGC AAAGCACACC TTGAGAAACA GAGAACTCTT  
17501 GTGCTCTCGC TCTCTTGACA CTTCAGGTGC AAAACTTTTG TCCTAATGTC  
17551 GTTCTCAAAC TTACGCATGT GTGAGAATCA CTGTGAGAGC TTATTGAAAC  
17601 TGATTGCGGG ACCCCATACC TAGAGGGCCT GATTCTATAG GTCTGAGGTA  
17651 AGGCCAAGA ATTTGCATAT TTGCATTTCT TTTTCTTTTC CTTTCTTTTC  
17701 TTTTCTTTT TTTTGTAGAT GAAGTCTCAC CCGTGCAGCC AGACTGGAGT  
17751 GCAGTGGCAT GATCTCAGCT CACTGCAGCC TCTGCCTCCT GGGTTAAAGC  
17801 GATTCTCCCC ACACCCAGAG CCCGCTCCTG AGTAGCTGGG ATTACAGGTG  
17851 CCCGCCACCA TGAAGTACTA ACGTTGTAT TTTTAGTAGA GACGGGGGT  
17901 TCACCATGTT GGCCAGGCTG GTCTCAAACCT CCTGACCTCA GGTGATCCAC  
17951 TCACCTCAGC CTCCCAAGGT CTTGGGATTA CTGGTGTGAG CCACCGCGTG  
18001 CGGCCAGAA TTGCATTTCT AACAGTCCC AGGTGATGCT GATGCTGTGG  
18051 GTCCAGGGAC ACACCTTGAG AACAGCTTGT TACTCAGGCG ATATGTGGAC  
18101 AGTAGCGTCA TCTTCACCTG GGAGCTTCCT GCAGCATCTC AGGCCTTGCC  
18151 CTACACCTAC CAGATCAGAA TCTGCATTTT AACTCAATCC CCGCGTGATT  
18201 CTCATGCACC TGGAAGTTTG AGAAATATGA CCTTAGAGGA GCCGGAATGT  
18251 GAAACCACTG GAGGCAGAGA TAGATGGAGA ATATCTCTTC TTCTACGGGA  
18301 TACTAAAGAT GCAACAAAAA GGGCTGACTC TCTGGTGTG CACCCAGGTG  
18351 GGGCTGATGA CCGAAAAGAG GCCAGATGTG GACAGAGGAC TCTTCCCTGA  
18401 GGGGAAGGCAG AGAGAACTTA GGAAAATCTG AAGAAAGGAG GTGGCTTCAG  
18451 AGGAAAGGCA TTCATCTGGG CCATAAAACA GTGGAGAAGG TATCCTGCTG  
18501 AGAGCACAGG GGTGGGGAGG GGGTGCCCTG GAGCTGAAGT CTTCAAGTGGG  
18551 GGGACAGTGA TAGGTGAACA CACATGTGAA TAAACAGTTT GCTAAGCAGC  
18601 TGCGAGGGCT GGCCAAGGTG AGAAAGCATC CGTCTGCAGA GGCCTCAATA  
18651 AGGCCAGTGT GTTGACTTTG TCCTGCAGTG CTCAGCAGTG GAAAAACCA  
18701 ACAGCCACGC AGGGAGAGGG AAGGAGCCAC GATGGGCACG GGTACTGGG  
18751 GCCAGGGCTT GACTGGTAGG TGGACACAGC TGAAGGCCCA GGTGTGTGG  
18801 GAACAGAGCG CAGAAGCAAT AGATTCTCT TGAAGATCCT TGGGCTGTTA  
18851 ACCTTTTTTA AATTTAAGAG AGGTTGTGTG GCGGGGAGG GAGGAAGGAA

FIGURE 3, page 6 of 14

18901	AATCCTTCAG	AAGACATAGA	CTTACTCTGT	TTCTTCCATC	ATATGTGAAT
18951	GCAATATGAAT	AGCCAAAAGG	TGAATAAAAC	ACATGTTCCC	AGGTGGCCAG
19001	TGAGACCTAG	GTTGCAAGAT	GGTGGGGTGT	GTGTGAGGCC	GGGGAGTGCT
19051	GCGAGCCCCG	GAATTCTCTA	GCCTTAGTCC	CCCGCCACAT	AGCTAAGAAG
19101	TGAGGGAGGA	GGTGAGAAGG	AGTCACTGCC	CAGCCTCACT	TCCGGTGGAG
19151	TACCCTGTCT	CCTTGTCAGT	TCTGTCTCTG	GGGACAGTTG	CCTGCTTTCA
19201	CCTCTCCCTC	CATCCCTCT	TCTCTCACAG	GGAAAAATTC	ACCTTAATAT
19251	TGGAAGTTCC	TCTCCTAGCA	AAGTCCTTCT	CAGGCACCCA	CAGGCAAAAA
19301	GGAAACTAAG	CAGAGTTAGG	GCTTCCAGGC	CTAGCCAACT	ACACGACTCT
19351	CCTCTTGCTT	CCCTAAGAAC	CAGCGCAAGG	GGCAGCGTGG	GTTCCAGCAT
19401	AGATGGACCT	GTGTTGGAAT	CTCTGCACGT	GCTGTGCTGA	CCCTGGCTAG
19451	CCATTGACCT	CTCTGAGCCC	TTGTTTCCTT	TCCACTAGGC	TCTCTGAGGG
19501	CAGGGGCCAT	GTCTTTTTC	CTGCTCTGTC	TGCACTGAGC	ACTGTGCAGG
19551	GCACATAGGA	AGTTCCCAT	AATGTTTGTG	GGATAAAGGA	AATAAAACCT
19601	TCTCTCTTCC	TGTCCCCCTT	GTGATGGCTT	TGCACAAGGC	ACTGTCCTTG
19651	GCCAGGTTTG	CTAGGCTAGT	GTGAGGATAA	ACCAGGTATA	TTACAAATTG
19701	GAGAAAATTT	CTCGTCTCTC	TTGGAAGAAG	GTGCTGTATC	ATGAAACAAG
19751	AATGTCTTGA	TTCCCTTCTA	TGCCAGGTAC	TGGGGAGAAA	CAGGTGCCTG
19801	ATAACCGTTG	ATCCAGGCAG	AAATAAGCAT	ACTCCTGCTT	CCCAAGGCCT
19851	GATGCTTCTC	TCCTTCCTCC	CTTCTCCTCC	CCTTCTCTTC	ACTCTTCTC
19901	TGCACACATG	GAAGAATGGC	TGCCAGGCAT	TGCCCATTTG	GAAAAGTACA
19951	GCTCAATGGA	TATGAATCAG	CTTGGGCAGG	CGAGAAATGA	TTCAAGTCTG
20001	ACCAAATCGA	TTTAGTTCAG	GTTGCCCGTT	CTGCATCTTT	TTTCCCTTGT
20051	AATTAAATGA	TGATTGGTCT	TGATGGTGGG	AAGGAAGAGA	CAGAATTAA
20101	TTTGTGTTGCC	TTTGTAGAAA	GCTGGGGACA	GCACAGATAA	GGGAAGATGT
20151	CTCCCATTTG	GCAAATAACT	GATGCGGAGG	TGGAGTGGCA	GTGGTGATGG
20201	GGATGCTGGT	GCCTTCAGGC	CTTCTGGGCC	GGGCAGTGCA	GCTGGTGGCA
20251	GACGGTTCGG	AACCTACCA	TGTTCCCATC	TGAAAACGTG	GGCTGATCAT
20301	GCCCACTCCT	GACCTTGCTC	CAGGGAGTAC	ACAAAGACGT	AAGCTTAATT
20351	AACCCACCAG	ACGTAGCTCT	TGAATCCCTG	GGCATAGTGC	CTGGGTATAG
20401	TTAGAGTTGG	GGAGAGGCAT	GGTCAGCAAA	ACAACCTCCC	TCATCTCTCT
20451	GTGTGCACTC	AGAGTCAAGC	TGGCTGTCTG	TGGTGGTGCT	GACTTCTCTT
20501	GCTGCAGATT	TCTCCAATAT	GTTTCTGCCC	TGCACGCATT	TGCCAAATCC
20551	CTTCGGTTTT	TTGTGTCTCG	TGGCAGCTTA	GCTCCTCCAG	CCCTTGGATG
20601	AAGAAGCGTG	GGAAGTCTTT	GCTTCTCTTC	CCTCCCGCAG	TGACATGCCA
20651	TGCCATGCCA	CTGCCTCTTC	ATCTGGTCTC	ATGACAGTCA	CTCATAAGCA
20701	CCCGCATGTA	CCCGGCCCTG	CACTAGCTCA	TGACAGCTGC	AGTCAATTGG
20751	GCCAGGTGCT	GTATCTCATC	CGGCTCCTC	AGCAACCCCTC	TGAGATACTG
20801	GTAATGTCCC	TGATGAAGAT	ATTTACTGAG	GCAGAAATGG	ACGCTCAGTG
20851	AAGCAAGGTG	CCTGATGTTA	TAGCAATGAG	CTATGAGTGG	CCAGAGGGAG
20901	GAGATAAGCT	CAGGCCGTAC	ACCAAAGCCC	ATGCTCCTTC	TAGTCAACCA
20951	CAGTGCCTCC	TATGGTGAAT	GAGTGAGTCA	GCAACCAAGA	CGCATGAGGC
21001	CTTCTTTTTG	GTGAGCCTTG	GCTGGGTGCT	GAGGCTTCAG	GTACAATCAT
21051	GGGTTGGAAG	AGCCCTCCTC	TCTCTCCACA	GTCTGGCACT	ATGACCCCTT
21101	CTGGTTATTA	ACAAGGCAAA	GAGAGAGAGG	GAAGAAAGCA	GGCAAATAAT
21151	GTGGGTTGCT	ATTCTTAGAG	ATTAGAATTT	CAGGAAGGAT	AAACACAGCG
21201	TTCTCTCCAG	AAGTATAAAT	AGGAAGACTT	CACACATGAC	TAGAACGAGA
21251	CATGTTTTAA	GTCTGTCTGAG	TAAGGCAGTG	ATGAAGTAGA	TTTCCCCAGA
21301	TTCACTCTCC	CTCCTCTGGG	TCCCCCAGGG	CCTTTACTTG	TGGCAACTTT
21351	CAGCTCAGGG	AGGGAGGAAA	GCCCCTTTCA	AAGCTTCAGA	TACTTCTCTA
21401	AGGTCAGTTT	CTGCTTAAAG	AAGGCCTTTA	CATTACTTCA	TCCCTTTGCC
21451	AAATTAAACT	GAAAGGAAAC	CTTTCAAGTG	TGATTGCCCTG	GCCCTTTTCT
21501	GTTCAATTTCT	CGTGGGTACG	CTTTCTAACT	TTCTTTCTTT	CTTCTTTCT
21551	TCAGGTGTTG	ACTTTAAGAT	GAAGACCATA	GAGGTAGACG	GCATCAAAGT
21601	GCGGATACAG	ATCTGGTGAG	CTGGGGAGGA	GGAGGAGGCA	GATGTAGGAG
21651	AAGAGGACTT	CTGGCTGCTC	CTTAGCTGCC	CCTGCCATGT	GTAAAATTCC
21701	TAGGCTTCAC	CTGGGATAAC	TGGCCACCTC	TCTGATGGAT	GGAAGCGAAG
21751	TCTCAGAAGC	CCATCTCTTC	CTATAAGCCT	TAATCTCCAA	CCTCTAAGAA
21801	ACTTTAGGGG	ATTGACTACA	AGCACCAAG	GGCAGGAATT	AGAAGGAACT
21851	GGCACACTAA	GCATTGTGAA	TTTATCTCAG	GATTAGGCTT	TGCCCTTGGG
21901	CTGTGCCACA	CTATGTTAAG	ATTGGAAGGA	AGGAGGCTAC	ACCCCCATC
21951	ATTTAGGGCG	AGACCCTGAG	AGAGTTCCTC	AGGATAGCAT	GATGAAGTTT
22001	CCACAGTAGC	AGAGGGTGCT	GCTGTGGCTC	TCTGCCTGAG	GTCTTGGAAG

FIGURE 3, page 7 of 14

22051 CACTGCCTTT GCCAGGTTT AGAGCTCCCT CTCAATTCCA CAGCAGTATG  
22101 GGCACCTGCT TCAGAGGTCC CATAGGGACT AGGGGTGTAG CAGCATCCCC  
22151 TGCCAACCTCC CATCCAACCA AATCTGGCCA CAGTGGCCAG ATTCCAGAGA  
22201 GCTGTCCAAG GCCTGTTCTG GCTGTGGCTT CTGGTTTCTG CCAGGAGGGC  
22251 AGTTGGCAGG AGGGGCCAAG GCCCTGCAGG CCTGGTCAGC ACCAGCACAG  
22301 ATGACCAGGC CTCTGACTGC AGATCCCTGT GGGGATCCAA GCATCCCTGG  
22351 TTTTTCACCC TTTAGCTCCC CAGTTTTTCC TACAAGGGGA CAGCTCTGCT  
22401 CTTCCCTTCC CCGTCTGTTT CCATGGTCCC TGCTCCTCTG AGGGACTGGC  
22451 TTTCTCCTGC AGGGACACTG CAGGGCAGGA GAGATACCAG ACCATCACAA  
22501 AGCAGTACTA TCGGCGGGCC CAGGTAAGCC ACCACATTGG GGGTTTCAAA  
22551 GTGGGAAGCT GCCACCCACA CTCCCAGCTC TGGGTATTTG AGATGTCTGT  
22601 GCCACGGATC CCCTAAATAC AGTTCGCTG CTTGGAGGAG CGCAGGGCGT  
22651 CTTTCAGCTG TTTACTGATC ATTTGTCCGT CCATTGTTCA TGGCCCACTC  
22701 ACTGCAGGCA GGCCCCTGCC CTCACCCCTG ACTTCCACCC TCCATCCTGG  
22751 GTCAAAGATC CAGGTCAAAG CATGTGGTGT CTTCTGCTG TAGAGAGTTC  
22801 TGTGATGGC CTGGGAGGCG GCAGTGGTGG GGTCTGAGAG AAGAGATATT  
22851 TCTGGATGCT GAGCAGGGAG AATGGGAGAG TGGGACCCAA CCTTTAAGTT  
22901 TCCACGGCCC CTTCTGGCCC CATGACTGCA CTCTCTCTGT GCATATCACA  
22951 TCTCTCTATT TCTCTCTCTC TCAGGGGATA TTTTGGTCT ATGACATTAG  
23001 CAGCGAGCGC TCTTACCAGC ACATCATGAA GTGGGTCAGT GACGTGGATG  
23051 AGGTAGGAGA TGCCACCTCA CTGCCGGGGT GTGGAGAGGG TGCCTCACCG  
23101 GGAAGGCAA GGGCAGGGCC AGATGGGAAG GCAAATGCTT CCAGGAAGCT  
23151 TTGCCTTCCA CAGCCCTGGA TGAAGACCTC TGGGTGAGTA AGACATGGGG  
23201 AAGAAACCGA AGCTGCCATG CCCTCACTCT CTATACCCTG CCAGGCCTCC  
23251 ACGGCTGTGT CTTTCCCGGA AATGAATTAG TTCCAAGTCT TCCCTGTGAG  
23301 CAGCTTCTTT CTTGAAATCT TGGGACCAGG TGGAGTTGCA AGATTGGGAT  
23351 CTAGTCTCTG CTCTGCACAA TAGCTGTGGA GCCTTGGGAA GCCATTGAA  
23401 TCCTCTGGGT CCCAGTTTCC TGTAGAATGA GGGCTGGACT TACATCCAAT  
23451 GTCCTTTCCA GCTCTGATAC CAGTGGTCTA ACCCAAGGAA GCACCAGTCT  
23501 TAGCCAGAGT GTCCTTCTACC CTAAGCTCTC CCCGTGATAC CCTTGAGGTC  
23551 AGCCATGGCA CTTGGGGGAG CCTGGCACCT GCATCCAGTC GGCCCAACCT  
23601 GTCCTTAGGG CTCTGGAATT GGTGGTGGGC TGGAGGCAGT GCAGACTCTG  
23651 TAGGGAAAT TGGGGGGGCA GGCAGCACTC ACTGGCTGTT CTGCCCATCC  
23701 TTTGTCCCTA GTACGCACCA GAAGGCGTCC AGAAGATCCT TATTGGGAAT  
23751 AAGGCTGATG AGGAGCAGAA ACGGCAGGTG GGAAGAGAGC AAGGGCAGCA  
23801 GGTAAGTGGA GGGAAAAGGC AAGTCCACCC CAGGTCTCTC GCTGGGCCTC  
23851 CAGGGCCAGT CCTGAGCGTG GGGACCTAGG GGTGTGTTCC CCAGTGGCAG  
23901 GTCCTCCAC ACCTCCCGAG CACCCCAAGG CCCTGGGGGA GTGGCCATCC  
23951 TCGGAAGGCT TGTTGTCTGG GTTTCAGGAC AGAAGCCCAG AGATTGCGGG  
24001 TCCATCCAGA AACAAAGACG TCATAGGCAG CAACTCTCCC AAGTCCAGGT  
24051 CCCCAAATGC AGGATTGCCC TCTGCTTAAG AGATCATCCC CGTGTTAGTA  
24101 ATGAAGGACT TCAAGTTGTC AACCTCTTCT CTGACAGCAT CCAGGCCTAG  
24151 CTGCCATGTT ACGGTCGAGA AATGATCTCC CATCCACCC AACACTCCCC  
24201 CACTCCTGTC CTTCTTACCC AGGAAAGAGC CAGGGAGGCA AATGAGGAGA  
24251 CAAAGAGCCA CAGCTGGAGA AGCCATGGGG GCAGAAAGGG TAGGAGGATG  
24301 ACGCTGAGGG AATGTCCAAG CATGCAGGGA GACCATCCTC CCAGAGAGCA  
24351 GAAAGAAATA TTGGTTATTT TTTTCTTCT TCTTCTTTT TTTTTTTTTT  
24401 TTTGAGATGG AGTCTCGCTC TGTCACCCAG GCTAGAGTGC AGTGGCGCCA  
24451 TCTCGGCTCA CTGCAACCTC TGCCTCCTGA GTTCAAGCAA TTCTTCTGCC  
24501 TCAGCCTCCC AAGTAGCTGA GATTACAGGT GCATGCCACC ACGCCTGGCT  
24551 AATTTTTTTT TATTTTTTAGT AGAGATGGGG TTTTGCCATA TTGGCCAGGC  
24601 CGGTCTCGAA CTCCTAACCT CAGGTGATCC ACCTGCCTCA GTCTCCCAA  
24651 GTGCTGGGAT TACAGGCGTG AGCCACTGTG CCCAGCCAAG ATTGGTATTT  
24701 CTGAGATAAG TTATCCACTC AGTCCGTGGA CCTCAAGAGT TTTCTCTCTC  
24751 CTTTTCAGTC AATAGCGTTC CATTAGTACT TAAATGAAA TTGATTGTTT  
24801 GGTATAAAAT ATAAGACATG GTCATTGACC AATTTGAAAG TAGAGGCAAA  
24851 GCCTACTAGG ATAGTATTTA TTGAGCACTC TATGTGTGGC ACTGTGCTAA  
24901 GGCAAGCGCT TTTAAGTGCA CGACCCCACT GAATCATCCC ACAACCATGG  
24951 ATGGGAGACA CACTCAGTCT CCTTTAACAG AAGATAAAGC TGGGGCTTAC  
25001 AGAGAATGTA CAACTTGTC AAGGTCACAC AGCTAGCCAT CAGTGGCAGT  
25051 GCTGCTATTC AGGCTGCGGA CTGTGGGACT CCAGAGCCCA TGTTTTTTAC  
25101 GAGGATGCCA TACTGCCACA ATGGATGGTG TCTTTATCTC CTGATATATG  
25151 ATTGTGTGTT GGGAGGCGTG GGGTGGCAGC TGGAAGAATG GAGAGGCATA

FIGURE 3, page 8 of 14

25201 TTTGTGGAGG ATCTTCCCCC ATTCTCTGCT ACCCTCTCTT GGAGCTCCCA  
25251 GTCCCATCTG AGAAATTATC TACTCTGAGA AATCGTCACA ACACAGCATG  
25301 GTTGTGAGTG CAGTGGCAGA AGCCTGTGCC TGGTTGTATG GGCCCCCTCCC  
25351 CTGCCTTACT GACTCTCTTT CAGAAATGTC CTTCTCTTGC AGCTGGCGAA  
25401 GGAGTATGGC ATGGAATTCT ATGAAACAAG TGCCTGCACC AACCTCAACA  
25451 TTAAAGAGGT GAGAGCCCTG GTGACCAGGC GCCCGCTCTC TCGGGCTGAG  
25501 TCCAGCAGAG GTGGGAGGAG GAGCCATAAG ATGGACCTTA TCCCTCAGGC  
25551 CGCTGCAGGG TTGCCAGGGG AGAGGAGGAG ACACTGGACT AACCTGTGCC  
25601 CTTTGGTTTC CAGTCATTCA CGCGTCTGAC AGAGCTGGTG CTGCAGGCC  
25651 ATAGGAAGGA GCTGGAAGGC CTCCGGATGC GTGCCAGCAA TGAGTTGGCA  
25701 CTGGCAGAGC TGGAGGAGGA GGAGGGCAAA CCCGAGGGCC CAGCGAACTC  
25751 TTCGAAAACC TGCTGGTGCT GAGTCCTGTG TGGGGCACCC CACACGACAC  
25801 CCCTCTTCCC TCAGGAGGCC CGTGGGCAGA CAGGGGAGCC GGGGCTTTGC  
25851 CCTGCTGCTG TCCTCTCGTG TGATGACCTT ATTGAGTATC AGTAGCCACT  
25901 ACTCCCCCTG CCGGCCCTG AGAGCGGCTC TGCTGTCATC TCAAGCAGCC  
25951 CCTGTCCCCA GCCCGTCCAC CCTGGAGTGG TCTTCTTCAG CCTGTTTCCC  
26001 CAGCCACAGG CTGCTACGA CCCCCACGAT GTGCCGCAAG CACTGTCTCA  
26051 CCATCCCGCA CCCACCAGAC AACAGCCAGG GCTGGAGTCC AGGCCACTTT  
26101 CAGCTGTCTC TTTCTCCGTG CATCGTGTCT CTTCTCTGCT TTTCTCTCT  
26151 TCCCCCACTT CTCTTTCTCT GACCCCTCCC CTCCGGTGCG TTTCTGTATCA  
26201 AAGCTCTCTA AACCCCGTCC CCCGTGTGTC CTGCTGTGTG CAGCTCGCTC  
26251 TTTCTTTTCT TCCTTAAGTA TCCAAGGGGA TGGACCCAGG CTGCTGGGGA  
26301 GGTTCACCCC TTGGATCCAG GAAGAACCCT CCACCCTGCC TCGTGGGTGG  
26351 GCCAAAGGCT ACAGGGTGCT TCTTCTCTT CCCCCACCCC CACTGTCCCT  
26401 CATGTGCCAT GGGCTGCCT CCCCAGTGAC CTGCGAAAGT GGAGCATCGA  
26451 GGTAGGAGGG AAACGGCAAC CAGGGAGTCC TCGAGCCTGG GGTGCCCCTA  
26501 CCTCTACCCA TTCCCCGACC AGAGCTTTGC CTTTGCTTGG CTGCCCCGCT  
26551 GCCTCTTTGG GGAACAGAGC TCAGAGGCAG GTGCTTCAGA GAAGGAAACA  
26601 AAATGAGGGG TGGCAGGGAT AAAAAGTCAC CTCCATTCTC TACCTCCCAT  
26651 GCAGCATGAA CACAATTTCT CTCCACCTGG CTCCCAAATT TAAAGATGTG  
26701 GACCAAGGCC TGTGGGTACT CCAGGGGCAA GGAGAGCCCT GGGGTGAGTG  
26751 ACAGTGTGAG GCCAACCATG CACTCCACAA AGGGGAGCAT TTGGAAATGA  
26801 AGGACTAGCT CCTATGTATC AGGTTAAGAG CAAGGGAGAG CTGGCCAGGG  
26851 ACAGCAGTTT GCACAGCAGA GGGGAATGTA GCAACAGCAG GGCTCCTAG  
26901 GCCCCATCTT CCATTTCTTA GGTAAGAAGA GCATTTCTCT AGACTCCCAG  
26951 GCGGAGGACT GAGCCTAGCC TTCAGCAACC AAGTTCTCTC TGGGACCCAA  
27001 AGTTTATGGG AGAAGGGCAA AGACTTCATG GGAAGAGAGA AGGAAGGCC  
27051 TGGGTAGAAA CGCTTGGTGC TGTCTCTTTT GGCCTTTAAG ACAAGCGCT  
27101 CATCTTGCCC TCTACCTCCT GATAGGCTTG AGGGTTTGCC AACCACTG  
27151 TGGCTACAGG TGGAGGGAAG AGGACTCCTT CCTCCAGAGT GCTATGTTCA  
27201 GGAAGTTTCT TTAACCCCAT ATGGCCCAAG AGTAGCTCGT AGGAGGCCCT  
27251 TTAAAGACGG AACAAAGTAAT TTACCAGTTC TACTGGGGTT CCTGCCCACC  
27301 GTCCCAAGGT GGGCGAGGCC TAGGAAGAGG GTCATTCTTA AGCCACACAT  
27351 TAGCTGCACT GCGTGGCTGC AGCCAAAACA AAGAACTGGG TGTGAGTAT  
27401 TCATCAACTA AGAACCAAAA TCCAGGGCAC TCATATGTGA AGGATAAGAA  
27451 CCTCACTTCC TTACTCCTCC AAAAAAGAAGT GGGGAAAGAA CCATCAAACC  
27501 TTTCTCTCTG ACTTACCAAA CCAGGAAAAC AGCAGGAGAG GGTGGCTCAG  
27551 GACTTAGGGA CAGGGTATAG CTTAGATGGT GGAAAGCAA GGAGAGCAGG  
27601 AAGTTGTAAA TCACTGGCTA ATGAGAAAAG GAGACAGCTA ACTCTAGGAT  
27651 GAAGCTGTGA CTAGGCTGGA GTTGCTTCCT TGAAGATGGG ACTCCTTGGG  
27701 TATCAAGACC TATGCCACAT CACACTGGGG CTAGGGAAGT AGGTGATGCC  
27751 AGCCCTCAAG TCTGTCTTCA GCCAGGGACT TGAGAAGTTA TATGGGCAG  
27801 TGGCTCCAAT CTGTGGACCA GTATTTGAGC TTTCCCTGAA GATCAGGCAG  
27851 GGTGCCATTC ATTGTCTTTC TCTCCTAGCC CCCTCAGGAA AGAAGGACTA  
27901 TATTTGTACT GTACCTTAGG GGTCTGTGAA GGGAAAACAT GGAATCAGGA  
27951 TTCTATAGAC TGATAGGCCC TATCCACAAG GGCCATGACT GGGAAAAGGT  
28001 ATGGGAGCAG AAGGAGAATT GGGATTTTAG GGTGCAGCTA CGCTCACCTT  
28051 AAACTTTGGG TGGCCTGGGG CATGTCTTGA GGCCAGACT GTTAACCAGG  
28101 CTCTGCTGGC CTGTTACTC GTCACCACCT CTGCACCTGC TGTCTTGAGA  
28151 CTCCATCCAG CCCCAGGCAC GCCACCTGCT CCTGAGCCTC CACTATCTCC  
28201 CTGTGACGGG TGAACCTCGT GTACTGTGTC TCGGGTCCAT ATATGAATTG  
28251 TGAGCAGGGT TCATCTATTT TAAACACAGA TGTTTACAAA ATAAAGATTA  
28301 TTTCAAACCA CCGGTGTGGC TGCCTGGATG AGTCCTTGGG GGTAGGTCTC

FIGURE 3, page 9 of 14

```

28351 ACTCAGACCC TGGCAGTGAT GTGGGAGGGA GAGAGGCAGT GCTGGTAGAA
28401 GCAGCTCCAG AAGCAAAGGC AACAGCAGTA GAGTGACCAC GGAAGCGGCA
28451 AACATTGTCT TCCCTTCTCT ACCTTCCCTA GTGCCACCTG CAGGGAGGCC
28501 CAAAGCAAAG CCCCGTTGCC CTGCATTGGG CTGGCACTGC AGAAATAAGA
28551 TGAAACACAG TTATCGAGAG GATGCTGAAC ATCTATGAGC AGGTTTAAA
28601 GCCAAGATGA GTCTCATCTG TTTGTGTGGG TCAGGAACGG GTCTTCCTGA
28651 AGGCATGAGG TGGGACTGGA TAATCTTTCA GATTGTGAT TGGATACCTC
28701 GGGGGAGCAG AGGCAGACTG GGATCTCAGG ACTGCAGGTA TTTCATACTT
28751 TGGGATATGG AATTGATGGA (SEQ ID NO:3)

```

#### FEATURES:

```

Start:      2044
Exon:       2044-2167
Intron:     2168-21554
Exon:       21555-21615
Intron:     21616-22462
Exon:       22463-22523
Intron:     22524-22974
Exon:       22975-23052
Intron:     23053-23711
Exon:       23712-23801
Intron:     23802-25392
Exon:       25393-25458
Intron:     25459-25613
Exon:       25614-25769
Stop:       25770

```

#### CHROMOSOME MAP POSITION:

Chromosome 14

#### ALLELIC VARIANTS (SNPs):

DNA	Position	Major	Minor	Domain
	206	-	T	Beyond ORF(5')
	4963	C	T	Intron
	8175	G	A	Intron
	10515	T	C	Intron
	13034	T	C	Intron
	13781	T	C	Intron
	14050	A	C	Intron
	14273	-	T	Intron
	17582	T	C	Intron
	17700	C	T	Intron
	18074	T	C A	Intron
	19328	G	T	Intron
	19570	A	G	Intron
	20892	C	T	Intron
	26465	G	A	Beyond ORF(3')
	26472	A	G	Beyond ORF(3')
	28071	C	T	Beyond ORF(3')
	28096	C	G	Beyond ORF(3')
	28403	A	G	Beyond ORF(3')
	28467	C	G	Beyond ORF(3')

Context:

DNA

Position

206

GCTCAAGATTGCACAGCTGGTGTAGTGGTGACACTGGGACTGGAACCCAAGTGTGCCTTAC  
TCCAGAGCCCTTGGCATGCACCTGAAACCCCATGTAAGCCCACTGTGGAGACGCGCACCT  
CGAAATAATGGAATCCACTACATCAGTTCCTTTAGCTTTCTGTGTAATCAGAGTAGCTAG  
CAGGCTCGGGATTTGCCCCCCCCGGC

[-, T]

TTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTTTTGCTCTTGTGCCCAGGCTGGAGTGCAAT  
GGCGCAATCTCGGCTCACCGCAACCTTCGCCTCTCAGGTTCAAGCAATTCTCCTGCCTCA  
GCCTCCCGAGTAGCTGGGATTACAGGCACCGGCCACCACGCCAGCTAATTTTTTTTATAT  
TTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTTTCCCTCTT  
ATTATAATTCAGACACTTAACCTGAAATATACCTTTTCAAATGAAGTAAATGGGCTTACC

4963

TATTAAGGGACTTGGGATTCTCCCTTATCTTGGGCGTGTTTTTTCAGCATTAACATAAACT  
TAAAGGAAAGAGTTGGATGGTCAAGAAAAGCTTTTTCCTTAAGTGATATGACAGTTTCT  
CAAGGAGGTAGAAGGGGAGCCAGGAGACAAATCAAGGAGCCAACGAAATGAGTGTACCT  
AAGTCATAGTCATTTCGCTTATTTTTTAAAAAATGCGTGTCTGTATGCCAGGCTCTGCACT  
GAGACCGAGAGATTCCAAGATGAATAATACCTACAGTCACTGTTCTCAAATTGTGCATTA  
[C, T]

CTAAACACATTACATGACCATGCTGGCCACTGATCGAGGCACCTTTCCCAGGGGCTTTT  
TTTGTGAATTAAAGAAAACAAGGTAATTCACCAGTTATTGCCAAGATAGTTGGCTTCTTG  
GCTCATGTGGATATCACCTAGGCCAGTACTTTTGTGATTACTGTGTACTCCACTTTAAC  
GGCCTGCGATCTCTAGAGAAGAACCCGCGCAGGGAGCAGTGAGAGGCCCTCCCTGGTAGAC  
TGAGACACTGACTGTCCCTCCCCCTATCCTTTTCGTCTTTCTGGCCAGCAGACCAGCAGG

8175

ATGCCAGGTGCCATGCTAAGATTGAGGACACAGTGGTGACCAAAACAGACAGAAACCAA  
GGAGCTGGCTTACATTCCAAGGGAGTGATAGGAAGCTGTGTTTCATTTAGTTTCTGCT  
CTAGTACCCCCCTTTCCCTGGCAGTGCCAGGGTCTGAGAAGGAAGAGTGAGGTGGTGAGG  
AGGTGTGAAGCAGTGGGGTGACCTGAGAGGAGAGGATGGGGTGGCTTTGCCTCAAGGCTT  
GGGCCCTGCTAGGTGTGCTCTGCCTCAGGCCTCTGTTTCTCCTCCTGACACAGGCACA  
[G, A]

ACTCGGCCTCCACCCCTTCCCCAAGGACATGACCTTGGGAAGGAACATATCTGAAGCCC  
GCGGAGGGTTTCCGCTGCTGTGCTATCTGTGCCACAGATCCGAGATGCACCCACAGCTGG  
GAGCACCGGTTCTCCCGCCTACCTGCACCTCCCTGGTTTCTGTTTCTCCTCCTCCTCCT  
TCCTTCTCCCCGCTCCCCAGACAGGCTGGTGATGAGCTTTATAACATGAAAGCTGATATT  
TGGCCATTATCCTTCTACCCTGATTGCCAGCTCTTCTCAGAGTGCCTTCTCTGTAATCC

10515

CTGGTGAAGGCTTTGAAGAGGAAGTGACATTTGAGTGGAGTCTTGAAGACTAGGCAGGAT  
TCTCCAGGGGCCCTGGGTGTGGGGGAAGCACACATCCTCTTCCCTGTAGGAGGTGCTGTG  
GAGAACACCTCCAGTGGGGCTGCTACTCTTCAGCCTTGCTGGGGCCAGCTGGAGTGGCCA  
CACCATGGTTCACACCAGCTGAAGTTCAAGAAGCCCCCTTGCCAGGAGATTGCTTTGCTGGC  
TCTGGGTGAGGGCAGGTGCATCTGGAAGCCCCCTTCTTTCTAAGATGTTTGTCTCCTGAGT  
[T, C]

TCTATGTCTAGTCTTTTCTTCCCTGAACCTTTTGTCTACCAGTCAGCACAGCCCTGCCTG  
AGAAGGAGGCTGGAGGAGTGAGTGGTCACTAGCCTGGTGGGTCTTGGCTGCCTCTGTGGT  
GCCGCTGGCCCTAAGTAGCAGGCTTAGGGAGGCGAGACCCAGTTCCAGGGGCTGCCAATG  
GGGAGCGAGATGGGGTGGCTGGAGCACACTGCACATGTACCAAGGCTCTAGGGAGGTCT  
GTGCACAAGGCAGTGGGAAAAGCAAGGGGAAGACCCAGCCTGGTCAACATGGTGAAACCC

13034

AGATTTGGGTGAGGACACAGCCAAACCATATCAGCTCCCGGGATCCCTGTGTGAATGGGG  
TCTTTTGGTGTGTTGAGGGCTGCACAGGGTGACCTCTTTAGAGGTGACCTCCTGCCACA  
ACCCACAGGAGGTGCACATGGCCACACATGCTGGTTTCTGCACTGGGAGGGGCTGGGG  
CACTCCTGGGACCTGTGCTTTGGTAACTGGAGCTGGCCTGGCCCTGGGGATTGGGTGTCTG  
CCTTGGGTTTCAGGTGTATTAGGTTGTTTCTCGTTGTGGAGTCTCATTACTAATGAAAAG  
[T, C]

TCAGGTCGCACTGCTGGTCTTTTGGGCTGTGGTTGATCCTGGTGATAACATTTGGCACCC  
AGAGGCAGCCCTGTTTCCACTGAAGCATGCGGAGCTTGGCTGGCAGGCAGGCAAGCTGGC  
AGCTGCCCTTAACCCATGAGGTGCTGGCCCGCTAGTAGGCACACCCTACCTGTGCCAGAA  
TTGAGGTTGTAGCCAGACTCCAGGAGCCATCTGGGCCCCACAGGGGGCGGCATTTCTCT  
TTTTGTTGAAACATTCCAGCCAAGTGCTGGCTTGGGCTTCATCTCTGTCCCCTCTCC

FIGURE 3, page 11 of 14



13781 CCCTGTGTTATGGGTTTTACACCTTATCTCACAATCTTAAAAAAAAAATTCTCTGAGAAAT  
CCTCTGTCACCCCACTTTACAGGTGAGGAACTGAGGCAAAGATAGGCTAACTGGCTTC  
CCCAACACCATGCAGTAATTAGTGATAAAGGCAGGGTTGGAACCAAACCTTGACCTCCA  
ATTGTGCTCTTAATGGCCAGGACACTCTGTGTCTTGAGCCACACTTCCTCCATGTTTCT  
AGGGCTTTCTAGGGAGGCAGACAGTGATGGGAAGGGGTGTTCTTTAGTGTGGATGTGCC  
[T, C]  
GCCTGCTCCTTTCTGTAAGCGTCACAGCACCTCCACTGCTGTACTGGGGAGGCACCAAGT  
TTTTCCCTGTTTGGCCACCCAAGGCGAGCTAGCTTAGGAGTCACGTGAGTGCTGGGTGTC  
TCGCCTGCTGCATCCCTCTATCCTGCCCCGCTGCCCCGGTGCCAGAGGAGGGCCCTGCCT  
GTCTTCCCAGTTCTCCAACAGCAGCGCTGTCCAGCACCTCGGGCTCCAGTTGTGGCCT  
GGCAGCTGTGGGGCAGACACCATAACAGACAGAGTCACAGCAGGAAGAGGATGGGGCCCA

14050 GGAAGGGGTGTTCTTTAGTGTGGATGTGCCCTGCCTGCTCCTTTCTGTAAGCGTCACAGC  
ACCTCCACTGCTGTACTGGGGAGGCACCAAGTTTTCCTGTTTGGCCACCCAAGGCGAG  
CTAGCTTAGGAGTCACGTGAGTGCTGGGTGTCTCGCCTGCTGCATCCCTCTATCCTGCCC  
CTGCCCCCGGTGCCAGAGGAGGGCCCTGCCTGTCTTCCCAGTTCTCCAACAGCAGCGCT  
GTCCAGCACCTCGGGCTCCAGTTGTGGCCTGGCAGCTGTGGGGCAGACACCATAACAG  
[A, C]  
CAGAGTCACAGCAGGAAGAGGATGGGGCCAGGGCTGCTGCCTCAGGCCATGGCTGCATG  
GCACCATCAGTTGATTGAGGAGCTTTCTTGCCAATGTCTGAGGCATCAGGTGGCAGGAC  
ACGTCTCCCTGCTCTTAAGCCTCAGGCATGCAGCCCTTCTTATGCTCTCTGGGGTGAGGG  
GGAGATCCCCCTCATGGAATTGCTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGGTCCT  
GCTCTGTCACTCAGGCTGGAGTGCAGCCTCAACCTCCAGACTCAAGTGATCCTCCTGCC

14273 TCTCCAACAGCAGCGCTGTCCAGCACCTCGGGCTCCAGTTGTGGCCTGGCAGCTGCTG  
GGGCAGACACCATAACAGACAGAGTCACAGCAGGAAGAGGATGGGGCCAGGGCTGCTGCC  
TCAGGCCATGGCTGCATGGCACCATCAGTTGATTGAGGAGCTTTCTTGCCAATGTCTGA  
GGCATCAGGTGGCAGGACACGTCTCCCTGCTCTTAAGCCTCAGGCATGCAGCCCTTCTTA  
TGCTCTCTGGGGTGAGGGGAGATCCCCCTCATGGAATTGCTTTTTTTTTTTTTTTTTTT  
[-, T]  
TTTTGAGACAGGGTCCTGCTCTGTCACTCAGGCTGGAGTGCAGCCTCAACCTCCAGACT  
CAAGTGATCCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACCACAGGTGGACACCATCACA  
CCTGGGTTTTTTTTGTTTTTGTTTTTTGTCTTAGAGATGGGGTCTCACTTTCTTGCTC  
AGTCTGGTCTCGAACTCCTGGGCGCAAGCAGTCTCCACCTCGTCTTCCAAAGTGTTT  
GGATTACAGGTGTGAGCACTGTGCTTGGCCTTTTATTTATTTAGAATTGTTTTTGAA

17582 GGATGTTTCTTCCATGACATATATAGCTCTTGAAACTACTTCTATCTAATATCACCCACA  
GTGCTGTTAAAAATACAGATTTCTGGGCCTCACCTCAAATATGATTAGTAGGTCTAG  
GCACGTCAAGGTCAATTGTTTTGTCTTTGTTTAAAGTCACCCAGGTGATTCTAAAGCCG  
AAGCTCTGCAAAGCACACCTTGAGAAACAGAGAACTCTTGCTCTCGCTCTCTTGACAC  
TTCAGGTGCAAACTTTTGTCTTAATGTCGTTCTCAAACCTACGCATGTGTGAGAATCAC  
[T, C]  
GTGAGAGCTTATTGAAACTGATTGCGGGACCCCATACCTAGAGGGCCTGATTCTATAGGT  
CTGAGGTAAGGCCCAAGAATTTGCATATTTGCATTTCTGTTTTCTTTCTTTCTTTCTT  
TTTTTTTTTTTTTGTAGATGAAGTCTCACCTGTGCGCCAGACTGGAGTGCAGTGGCATGA  
TCTCAGCTCACTGCAGCCTCTGCCTCCTGGGTAAAGCGATTCTCCCCACACCCAGACC  
CGTCTCTGAGTAGCTGGGATTACAGGTGCCCCGCCACCATGACTAGCTAACGTTTGTATTT

17700 AGGCACGTCAAGGTCAATTGTTTTGTCTTTGTTTAAAGTCACCCAGGTGATTCTAAAGC  
CGAAGCTCTGCAAAGCACACCTTGAGAAACAGAGAACTCTTGCTCTCTCGCTCTCTTGAC  
ACTTCAGGTGCAAACTTTTGTCTTAATGTCGTTCTCAAACCTACGCATGTGTGAGAATC  
ACTGTGAGAGCTTATTGAAACTGATTGCGGGACCCCATACCTAGAGGGCCTGATTCTATA  
GGTCTGAGGTAAGGCCCAAGAATTTGCATATTTGCATTTCTGTTTTCTTTCTTTCTTTT  
[C, T]  
TTTTTTTTTTTTTTTGTAGATGAAGTCTCACCTGTGCGCCAGACTGGAGTGCAGTGGCAT  
GATCTCAGCTCACTGCAGCCTCTGCCTCCTGGGTAAAGCGATTCTCCCCACACCCAGAC  
CCCGCTCCTGAGTAGCTGGGATTACAGGTGCCCGCCACCATGACTAGCTAACGTTTGTAT  
TTTTAGTAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCTCAAACCTCTGACCTCA  
GGTGATCCACTCACCTCAGCCTCCCAAGGTCTTGGGATTACTGGTGTGAGCCACCGCTG

18074 TGCAGCCTCTGCCTCCTGGGTAAAGCGATTCTCCCCACACCCAGACCCGCTCCTGAGT  
AGCTGGGATTACAGGTGCCCGCCACCATGACTAGCTAACGTTTGTATTTTAGTAGAGAC  
GGGGGTTTCCCATGTTGGCCAGGCTGGTCTCAAACCTCTGACCTCAGGTGATCCACTCA

FIGURE 3, page 12 of 14

CCTCAGCCTCCCAAGGTCTTGGGATTACTGGTGTGAGCCACCGCGTGCGGCCAGAATTTG  
CATTTTCTAACAAGTCCCAGGTGATGCTGATGCTGTGGGTCCAGGGACACACTTTGAGAAC  
[T, C, A]

GCTTGTACTCAGGCGATATGTGGACAGTAGCGTCATCTTCACCTGGGAGCTTCCTGCAG  
CATCTCAGGCCCTTGCCCTACACCTACCAGATCAGAATCTGCATTTTAACTCAATCCCCGC  
GTGATTCTCATGCACCTGGAAGTTTGAGAAATATGACCTTAGAGGAGCCGGAATGTGAAA  
CCACTGGAGGCAGAGATAGATGGAGAATATCTCTTCTCTCACGGATACTAAAGATGCAA  
CAAAAAGGGCTGACTCTCTGGGTGTGCACCCAGGTGGGGCTGATGACCGAAAAGAGGCCA

19328 TGTGTGTGAGGCCGGGAGTGTGCGAGCCCCGGAATTCCTCAGCCTTAGTCCCCCGCCA  
CATAGCTAAGAAGTGAGGGAGGAGGTGAGAAGGAGTCACTGCCAGCCTCACTTCCGGTG  
GAGTACCCTGTCTCCTTGTCTGTCAGTTCTGTCTCTGGGGACAGTTGCCTGTCTTCACCTCTCC  
CTCCATCCCCTCTTCTCTCACAGGGAAAAATTCACCTTAATATTGGAAGTTTCTCTCCTA  
GCAAAGTCTTCTCAGGCACCCACAGGCAAAAAGGAACTAAGCAGAGTTAGGGCTTCCA  
[G, T]

GCCTAGCCAACTACACGACTCTCCTTGTCTTCCCTAAGAACCAGCGCAAGGGGCAGCGT  
GGGTTCCAGCATAGATGGACCTGTGTTGGAATCTCTGCACGTGCTGTGCTGACCTGGCT  
AGCCATTGACCTCTCTGAGCCCTTGTCTTCTTCCACTAGGCTCTCTGAGGGCAGGGGCC  
ATGTCTTTTCACTGCTCTGTCTGCACTGAGCACTGTGCAGGGCACATAGGAAGTTCCCA  
TAAATGTTTGTGGGATAAAGGAAATAAAACCTTCTCTTCTCTCTGTCCTTGTGATGGC

19570 AAAGTCCTTCTCAGGCACCCACAGGCAAAAAGGAACTAAGCAGAGTTAGGGCTTCCAGG  
CCTAGCCAACTACACGACTCTCCTCTTGTCTTCCCTAAGAACCAGCGCAAGGGGCAGCGTG  
GGTTCAGCATAGATGGACCTGTGTTGGAATCTCTGCACGTGCTGTGCTGACCTGGCTA  
GCCATTGACCTCTCTGAGCCCTTGTCTTCTTCCACTAGGCTCTCTGAGGGCAGGGGCCA  
TGTCTTTTCACTGCTCTGTCTGCACTGAGCACTGTGCAGGGCACATAGGAAGTTCCCAT  
[A, G]

AATGTTTGTGGGATAAAGGAAATAAAACCTTCTCTCTTCTCTGTCCTTGTGATGGCTT  
TGCACAAGGCACTGTCTTGGCCAGGTTTGCTAGGCTAGTGTGAGGATAAACCAGGTATA  
TTACAAATTGGAGAAAATTTCTCGTTCTTCTTGAAGAAGGTGCTGTATCATGAAACAAG  
AATGTCTTGATTCCCTTCTATGCCAGGTACTGGGGAGAAACAGGTGCCTGATAACCGTTG  
ATCCAGGCAGAAATAAGCATACTCTGCTTCCCAAGGCCTGATGCTTCTCTCTCTCTCC

20892 CCTTGGATGAAGAAGCGTGGGAACCTTTTGCTTCTTCTTCCCTCCCGCAGTGACATGCCAT  
GCCATGCCACTGCCTCTTCATCTGGTCTATGACAGTCACTCATAAGCACCCGCATGTAC  
CCGGCCCTGCACTAGCTCATGACAGCTGCAGTCAATTGGGCCAGGTGCTGTATCTCATCC  
GGCCTCCTCAGCAACCCTCTGAGATACTGGTAATGTCCCTGATGAAGATATTTACTGAGG  
CAGAAATGGACGCTCAGTGAAGCAAGGTGCCTGATGTTATAGCAATGAGCTATGAGTGGC  
[C, T]

AGAGGGAGGAGATAAGCTCAGGCCTGACACCAAAGCCCATGCTCCTTCTAGTCAACCACA  
GTGCCTCCTATGGTGAATGAGTGAGTCAGCAACCAAGACGCATGAGGCCTTCTTTTTGGT  
GAGCCTTGGCTGGGTGCTGAGGCTTCAGGTACAATCATGGGTTGGAAGAGCCCTCCTCTC  
TCTCCACAGTCTGGCACTATGACCCCTTCTGGTTATTAACAAGGCAAAGAGAGAGAGGGA  
AGAAAGCAGGCAAAATAATGTGGGTTGCTATTCTTAGAGATTAGAATTTAGGAAGGATAA

26465 TTCTCTGACCCCTCCCCTCCGGTGCCTTTCGTATCAAAGCTCCTCAAACCCCGTCCCCCG  
TGTGTCTGTGTGTCAGCTCGCTCTTTCCTTCCCTTCCCTAAGCTATCCAAGGGGATGGA  
CCCAGGCTCGTGGGGAGGTTCCACCCTTGGATCCAGGAAGAACCCTCCACCCTGCCTCGT  
GGGTGGGCCAAAGGCTACAGGGTGTCTTCTCTCTTCCCCACCCCACTGTCCCTCATG  
TGCCATGGGCCTGCCTCCCCAGTGACCTGCGAAAGTGGAGCATCGAGGTAGGAGGGAAC  
[G, A]

GCAACCAGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCCATTCCCCGACCAGAGC  
TTTGCCCTTGTCTGGCTGCCCGCCTGCTCTTCTGGGGAACCTGAGCTCAGAGGCAGGTGCT  
TCAGAGAAGGAAACAAAATGAGGGGTGGCAGGGATAAAAAGTCACCTCCATTCTCTACCT  
CCCATGTCAGCATGAACACAATTTCTCTCCACCTGGCTCCCAAATTTAAAGATGTGGACCA  
AGGCCTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTGAGGCCAA

26472 ACCCCTCCCCTCCGGTGCCTTTCGTATCAAAGCTCCTCAAACCCCGTCCCCCGTGTGTCC  
TGCTGTGTGTCAGCTCGCTCTTTCCTTCCCTTCCCTAAGCTATCCAAGGGGATGGACCCAGGC  
TCGTGGGGAGGTTCCACCCTTGGATCCAGGAAGAACCCTCCACCCTGCCTCGTGGGTGGG  
CCAAAGGCTACAGGGTGTCTTCTCTCTTCCCCACCCCACTGTCCCTCATGTGCCATG  
GGCCTGCCTCCCCAGTGACCTGCGAAAGTGGAGCATCGAGGTAGGAGGGAACGGCAACC  
[A, G]

FIGURE 3, page 13 of 14

GGGAGTCTCGAGCCTGGGGCTGCCCTACCTCTACCCATTCCCCGACCAGAGCTTTGCC  
TTGCTTGGCTGCCCCGCTGCCCTTTTGGGGAAGTGAAGCTCAGAGGCAGGTGCTTCAGAGA  
AGGAAACAAAATGAGGGGTGGCAGGGATAAAAAGTCACCTCCATTCTCTACCTCCCATGC  
AGCATGAACACAATTTCTCTCCACCTGGCTCCCAAATTTAAAGATGTGGACCAAGGCCTG  
TGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCAGTGACACTGTCAGGCCAACCATGCA

28071 GCCAGGGACTTGAGAAGTTATATTGGGCAGTGGCTCCAATCTGTGGACCAGTATTTTCAGC  
TTTCCCTGAAGATCAGGCAGGGTGCCATTCAATTGTCTTTCTCTCCTAGCCCCCTCAGGAA  
AGAAGGACTATATTTGTACTGTACCCTAGGGGTCTGGAAGGGAACATGGAATCAGGA  
TTCTATAGACTGATAGGCCCTATCCACAAGGGCCATGACTGGGAAAAGGTATGGGAGCAG  
AAGGAGAATTGGGATTTTAGGGTGCAGCTACGCTCACCTAAACTTTTGGTGGCCTGGGG  
[C, T]  
ATGTCTTGAGGCCAGACTGTTAACCAGGCTCTGCTGGCCTGTTTACTCGTCACCACCTC  
TGCACCTGCTGTCTTGAGACTCCATCCAGCCCCAGGCACGCCACCTGCTCCTGAGCCTCC  
ACTATCTCCCTGTGACGGGTGAACCTTCGTGTACTGTGTCTCGGGTCCATATATGAATTGT  
GAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGATTATTTCAAACCAC  
CGGTGTGGCTGCCTGGATGAGTCTTGGGGTAGGTCTCACTCAGACCCTGGCAGTGATG

28096 GGCAGTGGCTCCAATCTGTGGACCAGTATTTTCAGCTTTCCCTGAAGATCAGGCAGGGTGC  
CATTCAATTGCTTTCTCTCCTAGCCCCCTCAGGAAAGAAGGACTATATTTGTACTGTACC  
CTAGGGGTCTGGAAGGGAACATGGAATCAGGATTCTATAGACTGATAGGCCCTATCC  
ACAAGGGCCATGACTGGGAAAAGGTATGGGAGCAGAAGGAGAATTGGGATTTTAGGGTGC  
AGCTACGCTCACCTAAACTTTTGGTGGCCTGGGGCATGTCTTGAGGCCAGACTGTTAA  
[C, G]  
CAGGCTCTGCTGGCCTGTTTACTCGTCACCACCTCTGCACCTGCTGTCTTGAGACTCCAT  
CCAGCCCCAGGCACGCCACCTGCTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAAC  
TCGTGTACTGTGTCTCGGGTCCATATATGAATTGTGAGCAGGGTTCATCTATTTTAAACA  
CAGATGTTTACAAAATAAAGATTATTTCAAACCACCGGTGTGGCTGCCTGGATGAGTCTT  
TGGGGTAGGTCTCACTCAGACCCTGGCAGTGATGTGGGAGGGAGAGAGGCAGTGCTGGT

28403 CTGCTGGCCTGTTTACTCGTCACCACCTCTGCACCTGCTGTCTTGAGACTCCATCCAGCC  
CCAGGCACGCCACCTGCTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGT  
ACTGTGTCTCGGGTCCATATATGAATTGTGAGCAGGGTTCATCTATTTTAAACACAGATG  
TTTACAAAATAAAGATTATTTCAAACCACCGGTGTGGCTGCCTGGATGAGTCTTGGGGG  
TAGGTCTCACTCAGACCCTGGCAGTGATGTGGGAGGGAGAGAGGCAGTGCTGGTAGAAGC  
[A, G]  
GCTCCAGAAGCAAAGGCAACAGCAGTAGAGTGACCACGGAAGCGGCAAACATTGTCTTCC  
CTTCTCTACCTTCCCTAGTGCCACCTGCAGGGAGGCCCAAAGCAAAGCCCCGTTGCCCTG  
CATTGGGCTGGCACTGCAGAAATAAGATGAAACACAGTTATCGAGAGGATGCTGAACATC  
TATGAGCAGGTTTTAAAGCCAAGATGAGTCTCATCTGTTTGTGTGGGTGAGGAACGGGTC  
TTCTGAAGGCATGAGGTGGGACTGGATAATCTTTCAGATTGTGATTGGATACCTCGGG

28467 GCACGCCACCTGCTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTACTG  
TGTCTCGGGTCCATATATGAATTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTA  
CAAAATAAAGATTATTTCAAACCACCGGTGTGGCTGCCTGGATGAGTCTTGGGGTAGG  
TCTCACTCAGACCCTGGCAGTGATGTGGGAGGGAGAGAGGCAGTGCTGGTAGAAGCAGCT  
CCAGAAGCAAAGGCAACAGCAGTAGAGTGACCACGGAAGCGGCAAACATTGTCTTCCCTT  
[C, G]  
TCTACCTTCCCTAGTGCCACCTGCAGGGAGGCCCAAAGCAAAGCCCCGTTGCCCTGCATT  
GGGCTGGCACTGCAGAAATAAGATGAAACACAGTTATCGAGAGGATGCTGAACATCTATG  
AGCAGGTTTTAAAGCCAAGATGAGTCTCATCTGTTTGTGTGGGTGAGGAACGGGTCTTCC  
TGAAGGCATGAGGTGGGACTGGATAATCTTTCAGATTGTGATTGGATACCTCGGGGAG  
CAGAGGCAGACTGGGATCTCAGGACTGCAGGTATTTCACTTTGGGATATGGAATTGAT